MEASURING LOW BIRTHWEIGHT AND PRETERM BIRTH IN RURAL NEPAL:

VALIDATING MATERNAL REPORTS, EXPLORING MATERNAL COMPREHENSION AND TESTING NEW METHODS FOR MULTIPLE IMPUTATION AND ADJUSTMENT

by
Karen T. Chang

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

August 2017

Abstract

Between 2011 and 2016, neonatal mortality in Nepal fell from 33 to 21 deaths per 1000 live births. With global resources allocated to scale-up interventions to meet the Sustainable Development Goals, tracking coverage of interventions and estimating the population in need are essential. Maternal reports collected as part of national surveys are often the only source of child health information available to generate nationally-representative coverage indicators in low-income countries. This research examines the validity of indicators calculated from maternal reports and explores possible strategies to improve these estimates in data collection and in analysis.

Nested in a large community-based randomized trial in rural Nepal, this thesis aims to address the three following research questions:

1. Are the low birthweight (LBW) and preterm birth indicators valid when calculated from maternal reports?
2. What are possible reasons for poor validity in these indicators?
3. How valid are new methods to adjust LBW calculated from recall data?

We interviewed mothers from the trial one to 24 months after birth about their child's health in the first week of life and compared reports to data collected prospectively in the trial to assess validity of their responses. We conducted focus groups with study staff and in-depth interviews with mothers who provided discordant responses to explore maternal understanding of questions. We then applied previously developed methods to multiply impute and adjust the LBW indicator to our dataset in a validation exercise.

Indicators generated from maternal reports underestimated the burden of LBW and preterm births in this study setting. The LBW indicator using reported birthweight in grams had low individual-level accuracy (AUC 0.69, 95% CI 0.67 to 0.72) and high population-level bias (IF 0.62, 95% CI 0.52 to 0.72). LBW using reported birth size and the preterm birth indicator had lower individual-level accuracy and higher population-level bias up to 24 months following birth. Challenges related to translations of questions and possible cultural-specific perceptions about birth size may have contributed to the poor validity of these indicators. Visual aids of newborns of varying sizes may help to scale relative birth size questions and facilitate more accurate maternal reports in different settings. In an analysis where patterns in missingness and heaping were simulated, new methods to multiply impute missing birthweights and adjust the LBW indicator performed better than previous methods.

Committee of Readers

Thesis Advisor:

Joanne Katz, ScD
Professor, Department of International Health

Committee Members:

Scott L. Zeger, PhD
Professor, Department of Biostatistics

Melinda K. Munos, PhD
Assistant Professor, Department of International Health

Li Liu, PhD
Assistant Professor, Department of Population, Family and Reproductive Health

Alternate Committee Members:

Luke Mullany, PhD
Professor, Department of International Health

Deanna L. Kerrigan, PhD
Associate Professor, Department of Health, Behavior and Society

Acknowledgements

This study was supported by the National Institute for Child Health and Development (HD060712) and the 'Bill and Melinda Gates Foundation' (OPP1084399, OPP1084442).

I am so thankful to the many people who have provided advice and support throughout this thesis research and my doctoral training. I would like to thank Dr. Luke Mullany for allowing me to nest this research in the Nepal Oil Massage Study and for his guidance throughout the project. To Drs. Melinda Munos and Li Liu, thank for the support and encouragement throughout this process as well as providing invaluable input on my proposal and dissertation. I would like to thank Drs. Scott Zeger and Deanna Kerrigan for serving as a member and alternate member on my committee of readers. And thank you to Cristina Salazar, Karla McCarthy, Brenda Casey, and Carol Buckley for all the help behind the scenes throughout the past four years.

I am so grateful to the NNIPS team for welcoming me to Nepal and for their direction throughout our trainings and data collection. Thank you to the data management team both in Baltimore and Nepal for taking such great care of the data from this project in addition to the many other projects they managed. Thank you to all the field staff who were so committed to their work on this study. And a special thank you to the women and their children who participated in this study.

To Dr. Joanne Katz, I am so appreciative of your mentorship over all these years. I admire your passion for research and your dedication to your students. And I am so lucky to have had the opportunity to learn from you and work under your guidance. Thank you to all my classmates in W5031 for sharing in all the ups and downs throughout this process and for being there to answer my many questions about sample size, regression, and life. I look forward to crossing paths again as we move into the next phases of our professional lives.

And last but certainly not least, I am grateful to my family and Ryan who have been so supportive throughout this process. Thank you for your continuous encouragement, for having confidence in me when I didn't always have confidence in myself, and for giving me the kick in the butt when I've needed it.

Table of Contents

Abstract ii

Acknowledgements iv

List of Tables viii

List of Figures ix

Chapter 1: Introduction 1

1.1 Targets to reduce child and neonatal mortality 1
1.2 Measuring newborn health indicators 2
1.3 Improving Coverage Measurement for MNCH 3
1.4 Burden of low birthweight and preterm birth 4
1.5 Measuring low birthweight and preterm birth 6
1.6 Prior research to validate the LBW and preterm birth indicators 7
1.7 Reasons for possible poor validity of LBW and preterm birth indicators 8
1.8 Methods to impute missing birthweight and adjust the LBW indicator 9
1.9 Summary 12
1.10 Reference 13

Chapter 2: Methods 19

2.1 Objectives 19
2.2 Study Population 20
2.3 Parent trial 21
2.4 Substudy 22
2.5 Addition of a qualitative component 24
2.6 Ethical approval 27
2.7 References 28

Chapter 3: Validation of maternal reports for low birthweight and preterm birth indicators in rural Nepal 29

3.1 Abstract 29
3.2 Introduction 31
3.3 Methods 33
3.4 Results 37
3.5 Discussion 43
3.6 Conclusion 47
3.7 References 49
3.8 Tables and Figures 53

Chapter 4: Mothers' perception of questions assessing birth size and length of pregnancy in rural Nepal 73

4.1 Abstract 73
4.2 Introduction 75
4.3 Methods 76
4.4 Results 81
4.5 Discussion 88
4.6 Conclusion 91
4.7 References 93
4.8 Tables and Figures 96

Chapter 5: Validation of adjustment and imputation methods for estimation of low birthweight prevalence using a rural Nepal dataset 100
5.1 Abstract 100
5.2 Introduction 102
5.3 Methods 107
5.4 Results 111
5.5 Discussion 115
5.6 Conclusion 119
5.7 References 120
5.8 Tables and Figures 123

Chapter 6: Discussion 134
6.1 Summary of results 134
6.2 Strengths and limitations 136
6.3 Programmatic implications 138
6.4 Future research and next steps 140
6.5 References 141

Appendices 142
7.1 Focus group discussion guide 142
7.2 In-depth interview guide 146

Bibliography 149

Curriculum Vitae 157

List of Tables

Table 3.1: Birthweight, birth size, and gestational age birth timing questions administered 53

Table 3.2: Characteristics of selected mothers & children 55

Table 3.3: Summary of birthweight and low birthweight 58

Table 3.4: Summary of gestational age and preterm birth 59

Table 3.5: Birthweight cards by place of delivery 60

Table 3.6: Overall sensitivity, specificity, AUC, IF 61

Table 3.7: 2x2 tables for low birthweight and preterm birth indicators 62

Table 3.8: Sensitivity, specificity, AUC stratified by child sex, place of delivery, maternal education, maternal age, and parity 64

Table 3.9: Sensitivity, specificity, AUC for LBW using reported birthweight by recall group 68

Table 3.10: Sensitivity, specificity, AUC, IF for LBW using reported birth size by recall group 69

Table 3.11: Sensitivity, specificity, AUC, IF for preterm using reported length of pregnancy by recall group 70

Table 3.12: Correct classification of newborns as low birthweight 71

Table 3.13: Correct classification of newborns as preterm birth 72

Table 4.1: Length of pregnancy and birth size administered in quantitative questionnaire 96

Table 4.2: Characteristics of mothers interviewed in qualitative follow-up 97

Table 5.1: Proportion of low birthweight, preterm birth and small-for-gestational-age in substudy and parent trial 124

Table 5.2: Normal mixture curves fitted to measured and reported birthweights 126

Table 5.3: Unadjusted and adjusted odds ratios (OR) for missing birthweight by socioeconomic and health risk factors among rural households in the 2011 Nepal DHS survey 129

Table 5.4: Missing birthweight for each significantly associated variable among rural households in the 2011 Nepal DHS dataset 131

Table 5.5: LBW point estimates (%) calculated using 6 methods in multiply imputed datasets (r=5) after removing birthweights at random and systematically 133

List of Figures

Figure 3.1: Flowchart for participant selection 54

Figure 3.2: Distributions of measured and reported birthweight 56

Figure 3.3: Distribution of gestational age 57

Figure 3.4: Inflation factors 63

Figure 3.5: Sensitivity and specificity of LBW and preterm births over recall time 67

Figure 4.1: Visual aids for birth size question 98

Figure 5.1: Flowchart for participant selection 123

Figure 5.2: Kernel density curve and one-component normal curves fitted to the measured and reported birthweight datasets 125

Figure 5.3: Kernel density curve, two-, and three-component normal curves fitted to the reported birthweight dataset 127

Figure 5.4: LBW estimates generated from six methods using measured birthweight dataset 128

Figure 5.5: Kernel density curve, one-, two-, and three-component normal curves fitted to a first imputation of missing birthweights after removing birthweights by rural wealth quintile 132

Chapter 1: Introduction

1.1 Targets to reduce child and neonatal mortality

Child mortality declined from about 10 million deaths in 2000 to an estimated 5.9 million deaths in 2015.[1] Despite these improvements in child survival in recent years, neonatal deaths declined at a much slower rate and now make up approximately 45% of all under-five deaths globally.[1] Thus, the first 28 days of life, especially the first week of life when around three-quarters of all neonatal deaths occur, is the time when a child is most vulnerable.[2,3] Unable to achieve the Millennium Development Goal (MDG) of a two-thirds reduction in child mortality between 1990 and 2015, the global health community has now set for itself a new target in the Sustainable Development Goals (SDGs) of reducing child mortality to no more than 25 per 1000 live births and neonatal mortality to no higher than 12 per 1000 live births in every country by 2030.[4] To track global progress towards achieving maternal, newborn and child health (MNCH) goals, indicators have been developed to align with the SDGs. Specific to neonatal mortality, the Every Newborn Action Plan aims to monitor coverage (the number who receive an intervention among the number within a population who need it), quality of these interventions received and equity across a population.[4–7]

Based on the 2016 Key Indicators Report from the Nepal Demographic and Health (DHS) Survey, child mortality fell to 39 deaths per 1000 live births from 54 deaths per 1000 live births in the 2011 Nepal DHS.[8] Neonatal mortality also decreased from 33 deaths per 1000 live births in both the 2006 and 2011 Nepal DHS surveys to 21 deaths per 1000 live births in the most recent five-year period.[8] Though substantial improvements have been seen in child and newborn survival in Nepal in recent years, mortality rates still remain above the SDG thresholds, requiring further efforts and interventions to combat threats to survival.

1.2 Measuring newborn health indicators

Tracking coverage of interventions and estimating the population in need of these interventions are challenging in countries where health records in facilities may rarely be available or reliable, particularly in settings where most births occur in the home.[9–11] In addition, those able to access care in facilities may differ from the broader population by health status, demographically, and socio-economically.[12,13] Maternally reported information, collected as part of national household surveys, like the Demographic and Health Surveys (DHS) and the Multiple Indicator Cluster Surveys (MICS), is often the only source of child health information available to generate nationally-representative coverage indicators in low-income countries.[10,11] The DHS and MICS are global household survey programs, the first supported by USAID and in operation since 1984 and the second developed by UNICEF in 1995.[10,14,15] Data collected as part of DHS and MICS surveys are frequently used to inform country health policies, to advocate for vulnerable groups, and to identify populations most in need of interventions.[10] As part of these surveys, mothers may be asked to remember events related to their child's health that might have taken place up to five years prior to administration of the survey.[10] More information about the accuracy of maternal reports and the degree of possible biases of these measurements across different settings is essential to improving interpretation and use of these data.[11] Given this reliance on survey data to monitor progress in meeting global targets for child health and the gaps in our knowledge of the accuracy of these indicators, a need has been identified to focus efforts to evaluate the validity of maternal reports of newborn health.

1.3 Improving Coverage Measurement for MNCH

A consortium representing international organizations, non-governmental organizations, and academic institutions as part of the Improving Coverage Measurement (ICM) Group housed within the Institute for International Programs (IIP) at the Johns Hopkins Bloomberg School of Public Health have taken on the mission to produce further evidence of the validity of MNCH coverage data, linking care-seeking indicators generated from household surveys to health service quality assessments, and facilitating the appropriate interpretation and use of coverage indicators.[16] Research of the ICM group is comprised of projects utilizing various study designs conducted in numerous settings, and each study design has its strengths and limitations. In many of these studies, validation of maternal and newborn health indicators comparing maternal reports to direct observations have been conducted in facility-based studies in Mozambique, Kenya and Mexico.[17–20] Use of direct observation as a reference, or “gold” standard, presents a more reliable and complete comparison than facility-based records. However, facility-based designs may be vulnerable to selection bias, and may often present the best case scenario in accuracy of maternal reports if factors associated with increased likelihood of facility births are the same as those associated with improved recall, like higher maternal education and wealthier households. To avoid possible selection bias, a previous study conducted in rural China used a population-based design; however, authors acknowledged that the reference standard of an existing electronic system may have contained some errors.[21] Using a population-based design, the study described in this thesis evaluates the accuracy of maternal reports of newborn health in a setting, where half of births take place in the home and where postnatal care for newborns is uncommon. What also makes this population distinct is its relatively high prevalence of low birthweight and preterm birth compared to the settings of these other studies. Prospective data collected as part of a larger community-based randomized trial served as the "gold" standard for this study. While the overall study collected data to validate a variety of newborn coverage indicators, this thesis will focus primarily on the low birthweight and preterm birth indicators. Other validation results will be published separately.

1.4 Burden of low birthweight and preterm birth

Low birthweight (LBW) has long been shown to be closely associated with a greater risk of neonatal death as well as cognitive and developmental impairment and long-term health problems in adulthood.[13,22,23] Historically, LBW, first described in the early 1900's as birthweight <2500 grams, was used to characterize premature newborns who had an increased risk of mortality and morbidity.[23,24] Weight measurements were widely available and gestational age data was often lacking, making identification of LBW babies more feasible.[25] However, upon recognition that not all LBW babies are born early and not all babies delivered early are of LBW, the conditions preterm and small for gestational age have become the favored constructs in recent decades.[25] Preterm birth is defined as a birth before 37 weeks gestational age, and small for gestational age (SGA), describes newborns with birthweights below the 10th centile of a birthweight-for-gestational age reference population.[26,27] Therefore, the term LBW captures both newborns born preterm and those who are small but not necessarily preterm; therefore, LBW and preterm are linked but not synonymous conditions.[27] While gestational age information is routinely documented in most developed countries, this remains challenging in many developing settings,[28] and as a result, LBW remains an important indicator of newborn health globally. In 2012, one of six global nutrition targets set by World Health Organization Member States was to reduce LBW incidence by 30% by 2025,[29] placing a renewed interest on how best to produce population-based estimates of LBW at the country level.

Approximately 20 million LBW infants are born annually, the vast majority from low and middle income countries (LMIC).[13] In South Asia 28% of infants are LBW, or one in every four births.[12] In the 2011 Nepal DHS report, of all children who were weighed at birth 12.4% were LBW and only 36.3% of children had a birthweight.[30] Each year, an estimated 15 million infants are born preterm, complications of which now constitute the leading cause of neonatal and under-five mortality.[1,31,28,32,22] More than 60% of preterm births occur in South Asia and Africa.[28,32] Preterm birth is associated with increased risk of cerebral palsy, vision and hearing impairment, and diminished learning abilities.[28,32–34] Given the increased vulnerability of infants born LBW and/or preterm, measuring and monitoring the burden of both LBW and preterm birth is needed to quantify the population in need of neonatal interventions in order to estimate coverage. Establishing denominators for these indicators is essential to the measurement of progress towards global targets aimed at improving child health outcomes.

1.5 Measuring low birthweight and preterm birth

One challenge to accurately tracking LBW is that more than half of children globally, and up to 69% of children in South Asia, are not weighed at birth.[13,12,29] Similarly, 64% of children were missing birthweights in the 2011 Nepal DHS, and 69% were missing among those selected from the central *terai* subregion, which includes Sarlahi District.[35] In South Asian settings, this is largely due to the fact that many births still occur at home and are thus not measured.[12,13,29] Additionally, often birth records in facilities in these settings are not available or not reliable.[13,12] Women who deliver in a facility are likely of a higher socioeconomic status compared to those who deliver at home.[13,12] Measuring gestational age is also problematic in low-resource settings; ultrasonography during the first trimester is not part of routine practice.[28] As a result, much of the data reported in these settings continue to rely on report of the first day of the last menstrual period, usually obtained late in pregnancy or at the time of delivery. However, inaccuracies in reports of the first day and fluctuations in menstrual cycles across women subject this measurement to errors.[28] Heaping, or the tendency of numerical measurements to be rounded to whole numbers, may also occur in both birth records from facilities and maternal reports, contributing to inaccuracies and misclassification of both LBW and preterm birth indicators.[36]

1.6 Prior research to validate the LBW and preterm birth indicators

Several studies in developed countries have demonstrated high accuracy of maternal recall of birthweight compared to hospital records.[37–41] Most such studies reported the intra-class correlation coefficient (ICC) comparing maternal reports to records of birthweight continuously rather than classifying infants dichotomously into either low or normal birthweight categories.[37–41] One study in Taiwan that asked mothers to recall their child's birthweight by phone reported mothers tended to overestimate birthweight and found maternal recall used to categorize children as LBW had low sensitivity (52%) and high specificity (95.3%).[42] In low-income settings, results have been heterogeneous. Findings from a facility-based study in Kenya of maternal reports of LBW at hospital discharge after birth and 13-15 months after birth compared to directly-observed deliveries found reports were accurate at both time points.[19,20] A study in Colombia assessing maternal recall of LBW five to 12 years after birth reported high specificity (95%) but moderate sensitivity (66%).[43] Similar to findings from Taiwan, a study in Uganda described mothers over-reporting birthweight four to 7 years after delivery[44] while another study in Brazil found that mothers of children who weighed less than 3500g at birth tended to overestimate birthweight while those with children weighing more tended to underestimate birthweight 11 years after delivery.[45]

Studies that assessed the relationship of birthweight and birth size within DHS datasets reported decreasing trends in birth size (very large, larger than average, average, smaller than average, very small) with decreasing birthweight.[46–49] However, when using maternal reports of birth size to calculate the indicator for LBW, sensitivity was low while specificity was high.[46–49] All such studies acknowledged that these analyses were limited by selection bias in that mothers who were able to report a birthweight in the DHS surveys were more likely to have delivered in a facility and be of higher socioeconomic status.

A number of studies have reported mothers could generally report gestational age accurately in developed countries.[37,39,40,50,51] One study conducted in the US Nurses' Health Studies population reported moderate sensitivity (68%) and high specificity (92%) using maternally reported gestational age to classify preterm birth.[38] Research to validate gestational age reports in developing countries is lacking since ultrasound during the first trimester is not routinely performed in many of these settings. Efforts to examine the validity of point-of-care ultrasound measurements performed by lower level staff as compared to readings by physicians are underway.

1.7 Reasons for possible poor validity of LBW and preterm birth indicators

Inaccuracies in maternal reports and misclassification of LBW and preterm births may not be able to distinguish between instances when participants did not accurately remember an event and when participants did not understand a question.[52–54] Feedback from data collectors themselves in conducting quantitative surveys can be useful to identify any additional probes that were provided to participants who had difficulty understanding a question and to understand the level of consistency in administering questionnaires across data collectors.[55,56] Prior studies have investigated respondents' comprehension of survey questions adapted from those in DHS and MICS questionnaires.[52–54] A key component of administering these questions across different settings is the issue of appropriate translation. In a study assessing the comprehension of questions in a Tanzania AIDS Indicator Survey, Yoder and Nyblade describe difficulties encountered with translation from English to Kiswahili, including problems with style and structure.[54] The authors encouraged the use of translations that are not literal, but rather, reflect the original intent of the question. Cognitive interviews may be a useful tool to determine participant comprehension following translation of a survey from English into a local language.[57–61] Another issue that may be relevant to improved interpretation of data from maternal reports is the influence of various community and regional factors within a societal context on mothers' perception of a child's size at birth.[62] This phenomenon may result in mothers across diverse settings providing different relative categorical birth size assessments for children of the same birthweight.[62] Identification of questions that may be difficult for mothers to understand and reasons for their misunderstanding may help in developing methods that could improve the quality of data collected in future surveys.

1.8 Methods to impute missing birthweight and adjust the LBW indicator

Given that over half of newborns globally are not weighed at birth[13,12,29] and mothers of newborns who were weighed at birth may not be able to produce a birthweight record or to report birthweight accurately at the time of the survey,[10,46,48,44,47] the need for methods to correct indicators generated from maternal reports is high. In an analysis of DHS surveys, Blanc and Wardlaw developed a method to adjust LBW estimates to account for heaping, the tendency of numerical birthweights to be rounded to multiples of 500g in maternal reports, and missing numerical birthweights.[13,63] To address heaping at the cutoff for LBW, after removing those who weighed exactly 2500g, the percentage of newborns who weighed less than 2500g for birthweights between 2000 and 2999g was calculated in 88 DHS surveys; this averaged to 25%. Based on these results, 25% of newborns who reportedly weighed exactly 2500g were reclassified as LBW. In these surveys, mothers are asked to report a relative birth size even if their child was not weighed at birth. To address missing birthweights, using available birth size information, the proportion of LBW newborns in each birth size category was calculated, multiplied by the overall proportion of births in each category, and summed to generate the overall number of LBW newborns. This correction is the current method used to adjust LBW estimates in MICS surveys,[63] and the adjusted estimate of LBW in Nepal in 2014 was 24.2%.[64]

A Working Group, comprised of members from UNICEF, WHO, the London School of Hygiene and Tropical Medicine and the Johns Hopkins Bloomberg School of Public Health, have recently developed new methods to adjust estimates of LBW calculated from maternally reported information. This group extended previous research describing the distribution of birthweight established across a variety of large datasets to an approach that aims to evaluate the quality of survey datasets, to identify a method using multiple imputation to handle missing birthweight, and to adjust LBW estimates accordingly. Wilcox *et al.* described birthweight as having a Gaussian distribution with a slight peak and an extended lower tail.[64] They identify two subpopulations that make up the larger distribution: a 'predominant' subpopulation with a Gaussian distribution that encompasses most birthweights and a 'residual' subpopulation made up primarily of LBW newborns.[64,65] Gage *et al.* asserted birthweight could be parametrically modeled as the sum of two Gaussian distributions and fitted both one- and two-component normal mixture models to birthweight data from different ethnic groups in New York state.[66] Charnigo *et al.* expanded on these results in his framework to propose the number of components in a normal mixture model of birthweight distribution may vary and help to identify heterogeneity in birthweight across ethnic populations.[67] Considering this prior research, the Working Group decided to apply similar normal mixture models in an approach to adjust LBW estimates and compare these to a crude estimate, the Blanc-Wardlaw method, and a kernel density estimate, which constructs a non-parametric curve using density estimates at each data point and a smoothing function. These methods were applied to publically available datasets of high data quality from the US and Mexico to first assess model fit and then applied to more than 200 DHS datasets to assess data quality.[68,69]

As part of the analysis of birthweight in each DHS dataset, the Working Group also identified a list of variables related to birthweight from a review of prior literature and in consideration of variables routinely collected in MICS and DHS surveys. A regression analysis was conducted in a number of DHS surveys to explore the association between birthweight and the related variables; the association was averaged across the surveys. From these results, the Working Group reduced the list to a set of variables that were overall significantly associated with birthweight. Multiple imputation was performed for missing birthweights predicted by an individual's observed values in the set of identified variables associated with birthweight. We investigate how these methods perform compared to existing methods for adjusting the LBW indicator. The process of developing this methodology for imputing missing birthweights and adjusting the LBW indicator will be described in detail in a separate publication.

1.9 Summary

Considerable improvements to child and newborn survival have been made in recent years in Nepal, however, mortality rates remain higher than the targets set as part of the SDGs. Accurate and timely measurements are required to monitor progress towards achieving these goals. As LBW and preterm newborns are at increased risk of mortality, valid estimates of these populations are necessary to identify those in need of neonatal interventions in calculations of coverage indicators. Our reliance on surveys for this information requires validation of questions administered in a variety of settings to understand possible limitations of this type of data and to identify methods that may facilitate improved data quality and interpretation.

1.10 Reference

1. Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. Lancet 2016;388(10063).

2. The World Health Organization (WHO). Postnatal care of the mother and newborn 2013. Geneva, Switzerland: 2013.

3. United Nations Children's Fund (UNICEF). The State of the World's Children 2016: A Fair Chance For Every Child. New York, NY: United Nations Children's Fund (UNICEF); 2016.

4. World Health Organization (WHO). Indicator and monitoring framework for the global strategy for women's, children's and adolescents' health (2016-2030). 2015.

5. World Health Organization (WHO), United Nations Children's Fund (UNICEF). Every newborn: an action plan to end prevantable deaths. 2014.

6. Marchant T, Bryce J, Victora C, et al. Improved measurement for mothers, newborns and children in the era of the Sustainable Development Goals. J Glob Health 2016;6(1):10506.

7. Grove J, Claeson M, Bryce J, et al. Maternal, newborn, and child health and the Sustainable Development Goals a call for sustained and improved measurement. Lancet 2015;386:1511–4.

8. Ministry of Health Nepal, New ERA, ICF. Nepal Demographic and Health Survey 2016: Key Indicators. Kathmandu, Nepal: 2017.

9. Moran AC, Kerber K, Sitrin D, et al. Measuring Coverage in MNCH: Indicators for Global Tracking of Newborn Care. PLoS Med 2013;10(5):1–9.

10. Hancioglu A, Arnold F. Measuring Coverage in MNCH: Tracking Progress in Health for Women and Children Using DHS and MICS Household Surveys. PLoS Med 2013;10(5).

11. Bryce J, Arnold F, Blanc A, et al. Measuring Coverage in MNCH: New Findings, New Strategies, and Recommendations for Action. PLoS Med 2013;10(5):1-9.

12. United Nations Children's Fund (UNICEF). Improving child nutrition. The achievable imperative for global progress. New York, NY: 2013.

13. United Nations Children's Fund and World Health Organization. Low Birthweight: Country, regional and global estimates. 2004.

14. United States Agency for International Development (USAID). The DHS Program. 2017 [cited 2017 Apr 29]; Available from: http://www.dhsprogram.com/

15. United Nations Children's Fund (UNICEF). Multiple Indicator Cluster Surveys. 2017 [cited 2017 Apr 29]; Available from: http://mics.unicef.org/

16. Improving Coverage Measurement in MNCH. 2017 [cited 2017 Apr 29]; Available from: http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/current-projects/improving-coverage-measurements-for-mnch/

17. Stanton CK, 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).

18. Blanc AK, Diaz C, Mccarthy KJ, Berdichevsky K. Measuring progress in maternal and newborn health care in Mexico: validating indicators of health system contact and quality of care. BMC Pregnancy Childbirth 2016;16:1–11.

19. Blanc AK, Warren C, Mccarthy KJ, Kimani J, Ndwiga C, Ramarao S. Assessing the validity of indicators of the quality of maternal and newborn health care in Kenya. J Glob Health 2016;6(1):1–13.

20. McCarthy KJ, Blanc AK, Warren CE, Kimani J, Mdawida B, Ndwidga C. Can surveys of women accurately track indicators of maternal and newborn care? A validity and reliability study in Kenya. J Glob Health 2016;6(2).

21. 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).

22. Unicef. Committing to Child Survival: A Promise Renewed. Unicef Progress Report 2014. 2014.

23. Wilcox AJ, Russell IT. Birthweight and Perinatal Mortality: II. On Weight-Specific Mortality. Int J Epidemiol 1983;12(3):319–25.

24. Salihu HM. The historic origin of Prematurity/Low birth weight: Lessons from the past. Am Acad of Pediatrics n.d.

25. Wilcox a J. On the importance--and the unimportance--of birthweight. Int J Epidemiol 2001;30(6):1233–41.

26. The World Health Organization (WHO). Physical Status: the Use and Interpretation of Anthropometry. Report of a WHO Expert Committee. Technical Report Series No. 854. 1995.

27. Lee AC, Katz J, Blencowe H, et al. National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. LancetGlobal Heal 2013;1(1):e26-36.

28. March of Dimes, PMNCHS, Save the Children, WHO. Born Too Soon: The Global Action Report on Preterm Birth. Eds CP Howson, MV Kinney, JE Lawn. Geneva, Switzerland: 2012.

29. WHO. Global Nutrition Targets 2025: Low Birth Weight Policy Brief. Geneva, Switzerland: 2014.

30. Ministry of Health and Population (MOHP) [Nepal], New ERA, ICF International Inc. Nepal Demographic and Health Survey 2011. Kathmandu, Nepal: 2012.

31. Howson CP, Kinney M V, McDougall L, Lawn JE. Born too soon: preterm birth matters. Reprod Health 2013;10 Suppl 1:S1.

32. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: A systematic analysis and implications. Lancet 2012;379(9832):2162–72.

33. Villar J, Abalos E, Carroli G, et al. Heterogeneity of perinatal outcomes in the preterm delivery syndrome. Obstet Gynecol 2004;104(1):78–87.

34. Alexander GR. "Prematurity at Birth: Determinants, Consequences, and Geographic Variation." Preterm Birth: Causes, Consequences, and Prevention. Washington D.C.: Institute of Medicine of the National Academies; 2007.

35. Central Bureau Of Statistics/Government of Nepal. National Population and Housing Census 2011. Kathmandu, Nepal: 2012.

36. Channon AAR, Padmadas SS, McDonald JW. Measuring birth weight in developing countries: Does the method of reporting in retrospective surveys matter? Matern Child Health J 2011;15(1):12–8.

37. Adegboye ARA, Heitmann BL. Accuracy and correlates of maternal recall of birthweight and gestational age. BJOG An Int J Obstet Gynaecol 2008;115(7):886–93.

38. Tomeo CA, Rich-Edwards JW, Michels KB, et al. Reproducibility and validity of maternal recall of pregnancy-related events. Epidemiology. 1999;10(6):774–7.

39. Sou SC, Chen WJ, Hsieh WS, Jeng SF. Severe obstetric complications and birth characteristics in preterm or term delivery were accurately recalled by mothers. J Clin Epidemiol 2006;59(4):429–35.

40. Bat-Erdene U, Metcalfe A, McDonald SW, et al. Validation of Canadian mothers' recall of events in labour and delivery with electronic health records. BMC Pregnancy Childbirth 2013;13(Suppl 1):S3.

41. O'Sullivan JJ, Pearce MS, Parker L. Parental recall of birth weight: How accurate

is it? Arch Dis Child 2000;82(3):202–3.

42. Li CY, Wei JN, Lu TH, Chuang LM, Sung FC. Mothers tended to overreport categorical infant birth weight of their children. J Clin Epidemiol 2006;59(12):1319–25.

43. Boeke CE, Marín C, Oliveros H, Mora-Plazas M, Agudelo-Cañas S, Villamor E. Validity of maternal birthweight recall among colombian children. Matern Child Health J 2012;16(4):753–9.

44. Lule S a., Webb EL, Ndibazza J, et al. Maternal recall of birthweight and birth size in Entebbe, Uganda. Trop Med Int Heal 2012;17(12):1465–9.

45. Araújo CL, Dutra CL, Hallal PC. Validity of maternal report on birth weight 11 years after delivery: the 1993 Pelotas Birth Cohort Study, Rio Grande do Sul State, Brazil. Cad Saude Publica 2007;23(10):2421–7.

46. Boerma JT, Weinstein KI, Rutstein SO, Sommerfelt AE. Data on birth weight in developing countries: Can surveys help? Bull World Health Organ 1996;74(2):209–16.

47. Mbuagbaw L, Gofin R. Can recall of birth size be used as a measure of birthweight in Cameroon? Paediatr Perinat Epidemiol 2010;24(4):383–9.

48. Eggleston E, Tsui AO, Fortney J. Assessing survey measures of infant birth weight and birth size in Ecuador. J Biosoc Sci 2000;32(3):373–82.

49. Islam MM. Can Maternal Recalled Birth Size be Used as a Proxy Measure of Birth Weight? An Evaluation Based on a Population Health Survey in Oman. Matern Child Health J 2014;18(6):1462–70.

50. Hakim RB, Tielsch JM, See LC. Agreement between maternal interview- and medical record-based gestational age. Am J Epidemiol 1992;136(5):566–73.

51. Poulsen G, Kurinczuk JJ, Wolke D, et al. Accurate reporting of expected delivery date by mothers 9 months after birth. J Clin Epidemiol 2011;64(12):1444–50.

52. Salasibew MM, Filteau S, Marchant T. Measurement of breastfeeding initiation: Ethiopian mothers' perception about survey questions assessing early initiation of breastfeeding. Int Breastfeed J 2014;9(13):1-8.

53. Yoder PS, Rosato M, Mahmud R, et al. Women's recall of delivery and neonatal care in Bangladesh and Malawi: A study of terms, concepts and survey questions. Calverton, Maryland, USA: 2010.

54. Yoder PS, Nyblade L. Comprehension of questions in the Tanzania AIDS Indicator Survey. DHS Qual Res Stud No 10 2004.

55. Salasibew MM, Dinsa G, Berhanu D, Filteau S, Marchant T. Measurement of delayed bathing and early initiation of breastfeeding: A cross-sectional survey exploring experiences of data collectors in Ethiopia. BMC Pediatr 2015;15(1).

56. Thapa S, Dhital M, Neupane S. Assessing the quality of survey data on adolescent sexuality by talking with the field staff. Contrib Nepalese Stud 2002;29(1):1–30.

57. Beatty P, Willis G. Research Synthesis: The Practice of Cognitive Interviewing. Public Opin Q 2007;71(2):287–311.

58. Giri J, Pokharel PR, Gyawali R, Bhattarai B. Translation and Validation of Modified Dental Anxiety Scale: The Nepali Version. Int Sch Res Not 2017;2017:1–5.

59. Huang KTL, Owino C, Vreeman RC, et al. Assessment of the face validity of two pain scales in Kenya: a validation study using cognitive interviewing. BMC Palliat Care 2012;11(5):1–9.

60. Harkness J. In pursuit of quality: Issues for cross-national survey research. Int J Soc Res Methodol 1999;2(2):125–40.

61. Jowell R. How comparative is comparitive research? Am Behav Sci 1998;42(2):1998.

62. Channon AAR. Can Mothers Judge the Size of Their Newborn? Assessing the Determinants of a Mother’s Perception of a Baby’s Size At Birth. J Biosoc Sci 2011;43(5):555–73.

63. Blanc AK, Wardlaw T. Monitoring low birth weight: An evaluation of international estimates and an updated estimation procedure. Bull World Health Organ 2005;83(3):178–85.

64. Central Bureau of Statistics. Nepal Multiple Indicator Cluster Survey 2014, Key Findings. Kathmandu, Nepal: Central Bureau of Statistics and UNICEF Nepal; 2014.

65. Wilcox AJ, Russellf IT. Birthweight and Perinatal Mortality: I. On the Frequency Distribution of Birthweight. Int J Epidemiol 1983;12(3):314–8.

66. Umbach DM, Wilcox AJ. A technique for measuring epidemiologically useful features of birthweight distributions. Stat Med 1996;15:1333–48.

67. Gage TB, Therriault G. Variability of Birth-Weight Distributions by Sex and Ethnicity: Analysis Using Mixture Models. Hum Biol 1998;70(3):517–34.

68. Charnigo R, Chesnut LW, Lobianco T, Kirby RS. Thinking outside the curve, part I: modeling birthweight distribution. BMC Pregnancy Childbirth 2010;10:37.

69. CDC National Center for Health Statistics. Birthweight and Gestation. 2017 [cited 2017 Apr 11]; Available from: https://www.cdc.gov/nchs/fastats/birthweight.htm

70. Secretaria de Salud Dirreccion General de Informacion en Salud. Databases of Live Birth Certificates. 2016 [cited 2017 Apr 11]; Available from: http://www.dgis.salud.gob.mx/contenidos/basesdedatos/std_nacimientos_gobmx.html

Chapter 2: Methods

2.1 Objectives

This thesis research has three main objectives: 1) to assess the accuracy of maternal reports of birthweight, birth size and length of pregnancy used in calculations for LBW and preterm birth indicators in rural Nepal, 2) to explore maternal understanding of administered questions about birth size and length of pregnancy and their perceived effect of the length of recall period on reporting accuracy in rural Nepal, 3) to validate new methods developed to multiply impute missing birthweights and adjust the LBW indicator using a rural Nepal birthweight dataset.

Specific aims

1. Validation of maternal reports for low birthweight and preterm birth indicators in rural Nepal
    a. To assess the validity of maternal reports of A) birthweight and B) birth size in correctly categorizing newborns as LBW, and C) length of pregnancy in classifying preterm births up to 24 months after delivery
    b. To determine whether the validity of maternal reports erodes with increasing time since delivery from one to 24 months after birth

2. Mothers' perception of questions assessing birth size and length of pregnancy in rural Nepal
    a. To describe maternal understanding of questions administered related to

their child's birth size and the length of pregnancy

b. To explore the perceived effect of time since delivery on accuracy of maternal reports among mothers in rural Nepal

c. To illustrate the potential utility of visual aids, like photographs, to help mothers gauge the size of their child at birth

3. Validation of adjustment and imputation methods for estimation of low birthweight prevalence using a rural Nepal dataset

    a. To apply previously developed methods to adjust the LBW indicator calculated from maternally reported birthweight

    b. To simulate missing birthweight patterns in the 2011 Nepal DHS dataset, multiply impute birthweights, and adjust the LBW estimate using previously developed methods

2.2 Study Population

Nepal is bordered by India and Tibet/China with 3 types of ecological zones: mountain, hill, and *terai* (or plains).[1] Of the estimated 28.5 million population in 2015,[2] about half live in the *terai* zones, which make up only 23% of the total land area in Nepal.[3] As of September 2015, Nepal is now divided into seven federal states (State 1, State 2, State 3, State 4, State 5, State 6, State 7).[1] The country's 75 administrative districts are further divided into smaller village development committees (VDC) in rural areas and municipalities in urban areas.[1] In 2011, 76% of households were involved in agriculture, and about 25% of the population lives below the poverty line.[3] Major ethnic/caste groups include Chhetri, Brahmins, Magar, Tharu, Tamang, and Newar.[3] In 2016, Nepal's neonatal mortality rate dropped to 21 deaths per 1000 live births after having remained relatively unchanged at 33 deaths per 1000 live births between 2006 and 2011.[1] The rural Sarlahi district in the *terai* region borders Bihar, India to the south. Of its 769,729 predominantly Hindu residents, 39% are less than 15 years of age.[4] Only 50.8% of the Sarlahi population five years and older are literate and literacy rates for men are about double those of women.[4]

2.3 Parent trial

This thesis project was nested within a randomized community-based trial in Nepal's *terai* region at the Nepal Nutrition Intervention Project, Sarlahi (NNIPS) field site. The site has been active since its first randomized control trial investigating the effect of vitamin A supplementation of mothers on child mortality in 1989.[5] From November 2010 to January 2017, the parent trial enrolled pregnant women and their newborns in 34 VDCs in Sarlahi District to investigate the impact of fully-body newborn massage with sunflower seed oil compared to mustard seed oil on neonatal morbidity and mortality. The trial was registered at ClinicalTrials.gov (NCT01177111). Newborn massage using mustard oil, often beginning immediately after birth, is widely practiced in this setting and across South Asia.[6] However, mustard oil may lead to a weakened skin barrier in newborns, especially in vulnerable preterm infants, and may ultimately increase the risk of morbidity and mortality.[6] In contrast, use of sunflower seed oil has been shown to positively impact neonatal skin and protect against infection compared to controls receiving no oil massage.[6] Locally-resident female project workers visited married women 15-35 years of age at home every 5 weeks to identify new pregnancies; pregnancies among women outside this age range were identified informally. All pregnant women agreeing to participate in the parent trial were asked to estimate the first day of their last menstrual period. Enrolled women were followed through delivery; study staff visited as soon as possible after delivery and through the first month (days 1, 3, 7, 10, 14, 21, and 28). At the first visit, workers recorded date and time of delivery, circumstances of labor and delivery, health status of mother and newborn, and the median of three measures of the baby's weight using a digital scale precise to 10g (Tanita BD-585). The date, time of birth and weight of the newborn were also provided to the mother/caretaker on a small 10 x 8 centimeter card. Subsequent visits focused on maternal report and directly observed aspects of newborn health. At each visit in the first month of life, study staff asked mothers to report on feeding and newborn care practices, and whether their newborn had any signs or symptoms indicating an illness in the days since their last visit. Staff also conducted examinations, including taking the newborn's temperature, measuring the newborn's rate of breaths per minute, and checking the newborn's cord stump and skin for signs of infection. For cases of illness or infection, staff referred mothers to nearby clinics or hospitals for care.

2.4 Substudy

Between April and September 2016, mother/child pairs that had participated in the parent trial were selected for one additional follow-up visit to ask mothers to report on events during labor and delivery, immediate newborn care, postnatal care, and cases of illness and care sought in the first 7 days of life. Mothers who had a singleton live birth and whose first visit by study staff had been conducted within 72 hours after birth were eligible. Given that a relatively low (i.e. <10%) proportion of newborns would have experienced illness or received postnatal care in the first week of life, we oversampled mother/child pairs with these characteristics. We defined illness as death or having two or more of the following signs in the same visit in the first week of life: difficulty sucking, difficulty breathing, stiffening of the back or convulsions, rapid breathing (a respiratory rate of 60 breaths per minute or faster), chest in-drawing, hyperthermia (100.4F or higher), hypothermia (lower than 95.9F), lethargy, or pus or redness at the base of the cord stump. Within each round of selection, we categorized newborns into four groups: those who experienced an illness, did not experience an illness, had a postnatal visit, and did not have a postnatal visit. We sampled all newborns who experienced an illness and/or had a postnatal care visit, and randomly sampled additional newborns without an illness and/or without a postnatal visit. Rather than aim to produce a representative sample of the larger community, our intent was to ensure we could evaluate the accuracy of maternal recall of more rare events such as care seeking for newborn illness. We conducted six rounds of selection from the parent trial and provided a list of eligible mother/child pairs in each round to our study staff trained in human subjects research and interviewing skills. In each round, additional mother/child pairs were selected as a buffer in case mothers were not at home at the time of a visit, had permanently moved, had died, or refused to participate. We aimed to interview approximately equal numbers of different mothers at each of seven follow-up time periods: 1, 3, 6, 9, 12, 18, or 24 months after birth. The first round selected mother/child pairs for the 18-month and 24-month recall group, the second round included those for the 6, 9, 12, and 24-month recall groups, the third round for the 1, 6, 9, 12, and 24-month recall groups, the fourth round for the 1, 3, 6, 12, and 24-month recall groups, the fifth ground for the 1, 3, 6, 12, and 24-month recall groups, and the sixth round for the 3, 6, and 12-month recall groups.

Study staff visited selected mothers and requested participation through an oral consent process in Nepali or Maithili, followed by collection of signature or thumbprint. If mothers did not want to sign or provide a thumbprint, a third-party, usually a family member or a neighbor, was asked to sign as a witness to her consent. Those agreeing were asked questions specific to this thesis related to birthweight, birth size, and length of pregnancy in addition to questions about newborn care practices, postnatal care, and morbidity and care seeking.

2.5 Addition of a qualitative component

During a presentation of preliminary results in April 2016, the addition of a qualitative explanatory component was suggested to supplement the quantitative data. A proposal was developed in June 2016 to explore maternal understanding of administered questions. We identified questions of interest from the quantitative form based on feedback from study staff during data collection supervision visits and on preliminary analyses of quantitative data that frequently generated discordant results comparing maternal report in the substudy to data collected in the parent trial. We focused on these questions in this qualitative analysis intended to complement and give further context to the interpretation of our quantitative findings. Logistical circumstances precluded data collection to continue until saturation; rather, we planned to complete a predefined number of FGD and IDIs prior to the start of data collection that would be logistically feasible. Two FGDs and 15 IDIs were planned at the start of qualitative data collection; however, because data collection for the parent trial was ending, we were unable to complete all planned IDIs.

From August to November 2016, we conducted focus group discussions (FGD) with the study staff, who had administered the quantitative form. A discussion guide was created to cover the following themes: reflection on experiences with administering the quantitative form, identification of questions mothers had difficulty answering, discussion of reasons difficulties were encountered, description of probes used for clarification, and suggestions for how questions could be improved for better understanding. Locally resident, female qualitative interviewers were fluent in Nepali and Maithili, from the same community as our study staff, and in non-supervisory roles in an attempt to allow study staff to more openly share their experiences working on the study. We conducted the FGDs in a private room at one of the field offices during the last month of data collection for the quantitative form.

Based on information from the FGDs, we narrowed our final list of quantitative questions of interest and identified and developed visual aids for use during in-depth interviews (IDI). We restricted IDI participation in this qualitative component to 9 Village Development Committees closest to our Sarlahi study office for logistical convenience. Mothers residing in these areas who responded discordantly to at least three of the questions of interest comparing data from the parent trial to their responses in the quantitative component of the substudy were eligible to participate in IDIs. Qualitative interviewers administered oral consent in Nepali or Maithili and obtained a signature or thumbprint for mothers who agreed to participate. IDIs were conducted in a private area in households of participants. An interview guide was created to cover the following topics: willingness to discuss labor and delivery and newborn health, attitudes about newborn health checks, views about whether time since birth affects mothers' ability to remember what happened, and reflection on questions that generated discordant responses and methods to improve the accuracy of maternal responses. FGD and IDI guides are included in **Appendices 1 and 2**. Visual aids, including dolls and photos of newborns of different sizes (**Chapter 4**), were developed based on suggestions from study staff during FGDs and used in administering questions related to birth size and demonstrations of newborn health examinations in IDIs. Analysis for this thesis focused on questions related to birth size and length of pregnancy and the effect of time since birth on mothers' memory. Results related to views about discussing labor and delivery and newborn health with others, attitudes about newborn health checks, and other questions that generated discordant responses will be reported elsewhere.

Discussion and interview guides were created in English and translated into Nepali and Maithili by local staff. Debrief sessions were conducted with qualitative interviewers following FGDs and IDIs to reflect on the quality of the discussion/interview, summarize content, edit questions for understanding, and discuss challenges. FGD and IDIs were audio recorded and transcribed from Maithili to Nepali by the interviewers. The Nepali transcripts were then sent to translators in Kathmandu for translation to English. For any additional clarifications that were needed in the English versions of transcripts, a native

Nepali speaker reviewed Nepali versions and re-translated sections. Recordings, transcripts, and translations were all de-identified.

2.6 Ethical approval

Both the parent trial and thesis project were approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board in Baltimore, USA. In Nepal, approval was received from the Tribhuvan University Institute of Medicine, Kathmandu for the parent trial and the Nepal Health Research Council, Kathmandu for the thesis project.

2.7 References

1. Ministry of Health Nepal, New ERA, ICF. Nepal Demographic and Health Survey 2016: Key Indicators. Kathmandu, Nepal: 2017.

2. The World Bank Group. The World Bank: Population, total: Nepal. 2017 [cited 2017 Apr 26]; Available from: http://data.worldbank.org/indicator/SP.POP.TOTL?locations=NP

3. Ministry of Health and Population (MOHP) [Nepal], New ERA, ICF International Inc. Nepal Demographic and Health Survey 2011. Kathmandu, Nepal: 2012.

4. Central Bureau Of Statistics/Government of Nepal. National Population and Housing Census 2011. Kathmandu, Nepal: 2012.

5. West KP, Pokhrel RP, Katz J, et al. Efficacy of vitamin A in reducing preschool child mortality in Nepal. Lancet 1991;338(8759):67–71.

6. Mullany LC, Darmstadt GL, Khatry SK, Tielsch JM. Traditional Massage of Newborns in Nepal: Implications for Trials of Improved Practice. J Trop Pediatr 2005;51(2):82–6.

Chapter 3: Validation of maternal reports for low birthweight and preterm birth indicators in rural Nepal

Karen T. Chang[1], Luke C. Mullany[1], Subarna K. Khatry[2], Steven C. LeClerq[1,2], Melinda K. Munos[1], Joanne Katz[1]

[1]Johns Hopkins University Bloomberg School of Public Health, Department of International Health, 601 N. Wolfe Street, Baltimore, MD, 21205, USA; [2]the Nepal Nutrition Intervention Project–Sarlahi, Lalitpur, Nepal

3.1 Abstract

Background

Tracking progress towards global newborn health targets depends largely on maternal reported data collected through large, nationally representative surveys. We evaluated the validity, across a range of recall period lengths (1 to 24 months post-delivery), of maternal report of birthweight, birth size and length of pregnancy.

Methods

We compared maternal reports to reference standards of birthweights measured within 72 hours of delivery and gestational age generated from reported first day of the last menstrual period (LMP) prospectively collected as part of a population-based study (n=1513). We calculated sensitivity, specificity, area under the receiver operating curve (AUC) as a measure of individual-level accuracy, and the inflation factor (IF) to quantify population-level bias for each indicator. We assessed if length of recall period modified accuracy by stratifying measurements across time bins and using a modified Poisson regression with robust error variance to estimate the relative risk (RR) of correctly classifying newborns as low birthweight (LBW) or preterm, adjusting for child sex, place of delivery, maternal age, maternal education, parity, and ethnicity.

Findings

The LBW indicator using maternally reported birthweight in grams had low individual-level accuracy (AUC 0.69, 95% CI 0.67 to 0.72) and high population-level bias (IF 0.62, 95% CI 0.52 to 0.72). LBW using maternally reported birth size and the preterm birth indicator had lower individual-level accuracy (AUC 0.58, 95% CI 0.56 to 0.60; AUC 0.56, 95% CI 0.53 to 0.58, respectively) and higher population-level bias (IF 0.28, 95% CI 0.22 to 0.34; IF 0.35, 95% CI 0.27 to 0.44, respectively) up to 24 months following birth. Length of recall time did not affect accuracy of LBW indicators. For the preterm birth indicator, accuracy did not change with length of recall up to 20 months after birth.

Interpretation

The use of maternal reports may bias indicators for LBW and preterm birth. In settings with high prevalence of LBW and preterm births, these indicators generated from maternal reports may be more vulnerable to misclassification. In populations where an important proportion of births occur at home or where weight is not routinely measured, mothers perhaps place less importance on remembering size at birth. Further work is needed to explore whether these conclusions on the validity of maternal reports hold in similar rural and low-income settings.

3.2 Introduction

Approximately 20 million low birth weight (LBW, <2500 grams) infants are born annually, the vast majority from low and middle income countries.[1] In South Asia 28% of infants are LBW, or one in every four births.[2] LBW is associated with increased mortality and morbidity, cognitive impairment, and long-term health complications in adulthood.[1,3] In 2012, one of six global nutrition targets set by World Health Organization Member States was to reduce LBW incidence by 30% by 2025.[4] Each year, an estimated 15 million infants are born preterm, complications of which now constitute the leading cause of neonatal and under-five mortality.[5–9] More than 60% of preterm births occur in South Asia and Africa.[7,8] Preterm birth is associated with increased risk of cerebral palsy, vision and hearing impairment, and diminished learning abilities.[7,8,10,11] Preterm and LBW are linked but not synonymous conditions. Measuring and monitoring both LBW and preterm birth indicators is essential to tracking progress towards global targets to improve newborn health outcomes.

One challenge to accurately tracking LBW is that more than half of children globally, and up to 69% of children in South Asia, are not weighed at birth. Despite substantial progress, many births still occur at home[1,2,4] and are thus not measured; in facilities weight is inconsistently measured, and records are often incomplete and/or unreliable, and women delivering in facilities may differ from the broader population in important ways (for example, by health status, demographically, socio-economically).[1,2] Measuring gestational age is also problematic in low-resource settings; routine ultrasound during the first trimester is not widely available, accessible or affordable.[7] This leads to reliance on reported date of last menstrual period (LMP), an error-prone measure given that collection is often late in pregnancy or at delivery, menstrual cycles vary in length, and non-negligible rates of lactational- or nutritional-amenorrhea exist in low resource settings.[7]

Thus, maternally-reported information collected through national surveys, like the Demographic and Health Surveys (DHS) and the Multiple Indicator Cluster Surveys (MICS), is often utilized to generate birth indicators in low-income countries.[12] Recall period are often long (i.e. up to five years prior to survey administration).[12] Improved understanding of the validity of maternal reports of newborn health is critical for monitoring global newborn health targets. We assessed the validity of postpartum reports of birthweight, birth size and length of pregnancy, by comparing maternal reports directly with data on birthweight and gestational age collected as part of a large community-randomized trial. By varying recall periods from 1 to 24 months, we examined if length of recall period modified indicator validity. We also assessed whether other maternal or newborn factors were associated with correct reporting of birthweight and preterm status.

3.3 Methods

Study setting

The study was conducted in Sarlahi district of Nepal, bordering Bihar, India to the south. Among approximately 800,000 predominantly Hindu residents, 40% are less than 15 years of age.[13] Government census data indicate that 15% of married women wed prior to age 15 and approximately 55.8% and 36.6% of males and females, respectively, five years and older can read and write.[13]

Parent trial

Between November 2010 and January 2017, a randomized community-based trial enrolled pregnant women and their babies in 34 Village Development Committees in the rural district of Sarlahi, Nepal to investigate the impact of full-body newborn massage with sunflower seed oil on newborn deaths and infections. The trial was registered at ClinicalTrials.gov (NCT01177111). Locally-resident female project workers visited married women 15-35 years of age at home every 5 weeks to identify new pregnancies; pregnancies among women outside this age range were identified informally. All pregnant women agreeing to participate in the parent trial were asked to estimate the first day of their last menstrual period. Enrolled women were followed through delivery; study staff visited as soon as possible after delivery and through the first month (days 1, 3, 7, 10, 14, 21, and 28). At the first visit, workers recorded date/time of delivery, circumstances of labor and delivery, health status of mother and newborn, and the median of three measures of baby's weight using a digital scale precise to 10g (Tanita BD-585). The date, time of birth and weight of the newborn were also provided to the mother/caretaker on a small 10 x 8 centimeter card. Subsequent visits focused on maternal report and directly observed aspects of newborn health.

Substudy

Between April and September 2016, mother/child pairs that had participated in the parent trial were selected for one additional follow-up visit to ask mothers to report on events during labor and delivery, immediate newborn care, postnatal care, and cases of illness and care sought in the first 7 days of life. Mothers who had a singleton live birth and whose first visit by study staff had been conducted within 72 hours after birth were eligible. Given that a relatively low (i.e. <10%) proportion of newborns would have experienced illness or received postnatal care in the first week of life, we oversampled mother/child pairs with these characteristics. We defined illness as death or having two or more of the following signs in the same visit in the first week of life: difficulty sucking, difficulty breathing, stiffening of the back or convulsions, rapid breathing (a respiratory rate of 60 breaths per minute or faster), chest in-drawing, hyperthermia (100.4F or higher), hypothermia (lower than 95.9F), lethargy, or pus or redness at the based of the cord stump. Within each of round of selection, we categorized newborns into four groups: those who experienced an illness, did not experience an illness, had a postnatal visit, and did not have a postnatal visit. We sampled all newborns, who experienced an illness and/or had a postnatal care visit, and randomly sampled additional newborns without an illness and/or without a postnatal visit. Rather than aim to produce a representative sample of the larger community, our intent was to ensure we could evaluate the accuracy of maternal recall of more rare events such as care seeking for newborn illness. We aimed to interview approximately equal numbers of different mothers at each of seven follow-up time periods: 1, 3, 6, 9, 12, 18, or 24 months after birth.

Study staff visited selected mothers and requested participation through an oral consent process in Nepali or Maithili, followed by collection of signature or thumbprint. Those agreeing were asked questions specific to this analysis (**Table 3.1**) in addition to questions about newborn care practices, morbidity and care seeking (results for those indicators will be published separately).

Ethical approval

The study was approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board in Baltimore, USA (parent trial and this substudy). In Nepal, approval was received from the Tribhuvan University Institute of Medicine, Kathmandu (parent trial) and the Nepal Health Research Council, Kathmandu (substudy).

Data analysis

Here we focus on assessing the validity of maternal reports of A) birthweight in grams and B) birth size in correctly categorizing newborns as LBW, and C) length of pregnancy in identifying preterm births. Maternal classification of LBW included a reported birthweight < 2500 grams (regardless of gestational age) or birth size of "small" or "very small." Maternal classification of preterm birth was defined as a reported time of delivery of "early" or "very early." We compared the percent of babies classified as LBW using these maternal reports of birthweight and birth size, to the proportion so classified using weight data collected using the digital scale (i.e. within 72 hours of birth). Similarly, we compared maternal classification of preterm birth with our “gold standard” estimate of pregnancy length estimated by calculating the difference in days between delivery date and reported first day of the LMP, collected through prospective surveillance as described above.

Sensitivity, specificity, area under the curve (AUC), and the inflation factor (IF) were calculated. AUC is the area under the receiver operating characteristic curve (plot of 1–specificity vs. sensitivity). An AUC of 1 represents a perfect test while a score of 0.5 suggests the test is no better than a random guess. The IF is the ratio of the survey-based prevalence of the indicator (given the estimated sensitivity and specificity) and the ‘true’ prevalence based on a gold standard, and thus quantifies the extent to which the indicator is over- or underestimated in a survey. The prevalence of an indicator in a population-based survey is calculated as Prevalence × (Sensitivity+Specificity-1)+(1-Specificity).[14] Bootstrapping with 1,000 replications was used to estimate standard errors and construct 95% CIs for the IF ratio.

We stratified sensitivity, specificity, AUC, and IF by child sex, birth location (facility vs. home), maternal education (any vs. none), maternal age (<20 vs. >=20 years) and parity (primiparous vs. multiparous). To examine whether accuracy of maternal reports erodes with longer recall periods, sensitivity, specificity, AUC, and IF were calculated within each of the seven bins of recall time. Comparable to other studies that have assessed validity of indicators, we defined high individual-level accuracy as AUC>0.70, and low population-level bias as 0.75<IF<1.25.[15–17]

Finally, to control for possible differences in the types of mothers interviewed across recall periods, we used a modified Poisson regression with robust variance estimate to estimate relative risk ratios (RR)[18] to assess the effect of time as a continuous variable on the proportion of newborns identified as LBW or preterm, controlling for child age, child sex, place of delivery, maternal age, maternal education, parity, and ethnicity. Stata version 14.0 (StataCorp, College Station, TX, USA) was used for the analyses.

3.4 Results

1528 mothers consented and were interviewed (**Figure 3.1**). After excluding 15 participants (birth assessment >72 hours after birth [n=3], twin delivery [n=1], repeat participation [n=11]), a total of 1513 mother/child pairs were included. Of these, 222 were enrolled in the one month recall group, 208 in the three month group, 206 in the six month group, 196 in the nine month group, 193 in the 12 month group, 289 in the 18 month group, and 199 in the 24 month group. The mean recall period was 10.7 months (**Table 3.2**). More than half of newborns were male (55.4%) and a majority of these births occurred in the home (53.8%). The mean age of mothers at follow up was 24.8 years; most had no schooling (68.4%) and had prior children (75.1%). Participants were nearly universally (96.2%) of Madhesi ethnic origin, frequently lacked a household latrine (71.2%), had electricity (80.3%), and owned some type of land (97.4%). Our sample was largely comparable to the parent trial sample; one difference was that the parent trial sample was more balanced by child sex (male=51.3%).

Mother/child pairs missing either the actual digital weight measures or the maternal assessment of weight were excluded. The former included deaths prior to measurement (n=14), parental refusal of weight measurement (n=1), and missing weight measurement (n=1); the latter exclusions included pairs where mothers reported the child was not weighed (n=21), was uncertain if child was weighed (n=6), or was weighed but could not respond (n=47). A further six mothers without LMP data were excluded. **Figure 3.2** provides the distributions of measured and reported birthweights, and **Figure 3.3** describes the distribution of calculated gestational age. Measured birthweights and calculated gestational age appear to be normally distributed with medians at 2750g (IQR=2460-3000g) and 39.6 weeks (IQR=38-40.9 weeks), respectively. Reported birthweights are heavily heaped, and have higher median (3000g) and a larger spread (IQR=2500-3500g) compared to measured birthweights. Results in **Table 3.3** are analogous with a greater mean and standard deviation in reported birthweights (mean=2886g, SD=608g) compared to measured birthweights (mean=2726g, SD=435g). A higher percentage of newborns were categorized as LBW when using the measured values (27.7%, 95% CI=25.4-30.0%) compared the reported values (17.1%, 95% CI=15.2-19.1%). This pattern is generally consistent when examining birthweight and percent LBW by sex and by birth size. Within both measured and reported values, mean birthweight increased and the percentage of newborns identified as LBW decreased with increasing birth size, though the trend was not statistically significant. Mean calculated gestational age was 39.3 weeks (SD=2.9 weeks) with 16.1% (95% CI=14.3-17.6%) of newborns categorized as preterm births, comparable by sex (**Table 3.4)**. Though not statistically significant, mean gestational age increased with increased reported length of pregnancy and the percentage of preterm births decreased with the exception of the late and very late groups. Compared to measured birthweight and gestational age in this study, the parent trial had a higher prevalence of LBW (29.9%) and lower prevalence of preterm births (15.4%).

Of the 1486 mothers who were asked if they had a card with a birthweight record, only 75 (9.4%) of those who delivered at home and 53 (7.7%) of those who delivered at a facility were able to produce one (**Table 3.5**). Of the total of 128 cards presented, 22 were from a facility and 106 were from the parent study. Of the 22 facility cards, all had reported delivering at a facility, and of the 106 birth cards distributed during the parent trial, 70.8% delivered in the home and 29.2% delivered in a facility. Comparing the percent of newborns that would be categorized as low birthweight based on birthweight measurements taken during the parent trial versus the birth cards produced at the follow up visit, fewer would be categorized as low birthweight for home births (17.3% vs. 20.0%) and more would be identified as low birthweight among facility births (26.4% vs. 22.6%). This observation is purely descriptive as no statistical test was performed with so few birth cards presented.

Sensitivity, specificity, AUC, and IF are presented in **Table 3.6** for the following indicators: A) LBW newborns based on maternally reported birthweight, B) LBW newborns based on maternally reported birth size, and C) preterm births based on maternally reported length of gestation (absolute numbers available in **Table 3.7**).

Sensitivity for all three indicators was low while specificity was high. LBW, estimated from maternal reports of birthweight, had low individual reporting accuracy (AUC=0.69, 95% CI=0.67-0.72) and high population-level bias (IF=0.62, 95%CI=0.52-0.72). Using reports of birth size and pregnancy length to estimate LBW and preterm birth prevalence, respectively, also resulted in low individual-level accuracy and high population-level bias.

To further investigate population-level bias, we plotted the values of the predicted survey prevalence of each of the three indicators across all possible prevalences of LBW and preterm births within our reference standard (**Figure 3.4**).[17] We would not expect most populations to have a prevalence of LBW and preterm newborns in the high ranges; therefore, this figure is for illustrative purposes. The gray dotted line represents perfect reporting accuracy with 100% sensitivity and specificity across all possible prevalences within our reference standard. The red, blue and green lines show the estimated survey-based prevalence and the differences in the predicted survey prevalence and the 'true' prevalence using the estimated levels of sensitivity and specificity for each indicator. All three indicators, with low sensitivity and high specificity, underestimate the survey-based prevalence in our study population, and would underestimate the survey-based prevalence to a greater degree in populations with higher prevalences of LBW and preterm births. In lower-prevalence populations, the bias would be lower. Assuming sensitivity and specificity remain the same as prevalence changes, survey-based estimates would underestimate the magnitude of changes in prevalence when looking at time trends or across countries. Stratified analyses by child sex, place of delivery, any maternal education, maternal age, and parity did not generally produce significantly different results (**Table 3.8**).

We observed no significant differences in the measures of accuracy by length of recall for any of the three indicators (**Figure 3.5** and **Tables 3.9-3.11**). These findings were consistent with our estimated RR for the proportions accurately categorized as LBW and preterm against recall length, controlling for child sex, place of delivery, maternal age, maternal education, parity, and ethnicity (**Tables 3.12-3.13**). In Model A, the estimated RR for the proportion of newborns correctly identified as LBW by maternally-reported birthweight negligibly decreased with increasing length of recall time and was statistically significant (RR=0.99, 95% CI=0.99-0.99, p-value=0.003), adjusting for other factors. In Model B, the estimated RR for the proportion of newborns correctly identified as LBW by maternally-reported birth size similarly decreased only slightly with increasing length of recall time but was not statistically significant (RR=0.99, 95% CI=0.99-1.00, p-value=0.09), controlling for other variables. For both models, mothers' reports were less likely to accurately identify newborns as LBW if their child was female, and were more likely to accurately categorize the child if the mother reported having had any education and if the mother had one or more children prior to the child of interest. Model C included an inflection point at 20 months in the time since birth. Up to 20 months after birth, the estimated RR for the proportion of newborns correctly identified as preterm by maternally reported length of gestation was not associated with time (RR=1.00, 95% CI=0.99-1.00, p-value=0.12), adjusting for other variables. After 20 months since birth, the estimated RR for the proportion of newborns correctly identified as preterm by maternally reported length of gestation improved slightly with increasing time since birth and was statistically significant (RR=1.02, 95% CI=1.01-1.04, p-value=0.001). Maternal age, parity, and ethnicity were also predictive of correct categorization of preterm birth.

3.5 Discussion

Accurate LBW and preterm estimates are necessary for assessing prevalence and denominators of indicators to assess coverage of interventions aimed at improving neonatal outcomes. We found that maternal reporting of birthweight, birth size, and length of gestation underestimates the true prevalence of LBW and preterm birth. LBW using maternally reported birthweight had low individual-level accuracy and high population-level bias up to 24 months following birth. Several studies in high-income settings have demonstrated high accuracy of maternal recall of birthweight compared to hospital records.[19–23] However, one study in Taiwan that asked mothers to recall their child's birthweight by phone reported mothers tended to overestimate birthweight and found maternal recall used to categorize children as LBW had low sensitivity (52%) and high specificity (95.3%).[24] This is consistent with our findings though children in Taiwan were much older at the time of follow-up. In low-income settings, results have been heterogeneous. Findings from a validation study in Kenya of maternal recall of LBW at hospital discharge after birth and 13-15 months after birth compared to directly-observed deliveries found high reporting accuracy and low population-level bias at both time points.[17,25] A study in Colombia assessing maternal recall for LBW 5-12 years after birth reported high specificity (95%) but low sensitivity (66%).[26] Similar to findings from Taiwan, a study in Uganda described mothers over-reporting birthweight 4-7 years after delivery[27] while another study in Brazil found that mothers of children who weighed less than 3500g at birth tended to overestimate birthweight while those with children weighing more tended to underestimate birthweight 11 years after delivery.[28] One possible reason for the lower sensitivity and higher degree of overestimating birthweight in our study compared to other studies may be partially explained by the relatively smaller size of newborns in this rural Nepal population. With more babies clustered around the cutoff of less than 2500g, we observed a high likelihood of misclassification in a population with high prevalence of LBW (27.7%).

We found the LBW indicator based on reported birth size had both low individual-level accuracy and high population-level bias. This indicator in our study had much lower sensitivity and higher specificity compared to that described in a study in Uganda (Sn=76%, 95%CI: 50–93% and Sp=70%, 95%CI: 65–75%).[27] Other studies that assessed the relationship of birthweight and birth size within DHS datasets found mean birthweight generally decreased with decreasing birth size, consistent with our findings.[29–32] However, when using maternal recall of birth size as an indicator for LBW, sensitivity was low while specificity was high.[29–32] All such studies acknowledged that these analyses were limited by selection bias in that mothers who were able to report a birthweight were more likely to have delivered in a facility and be of higher socioeconomic status. Our study results demonstrate similar findings to these prior studies in a population with more than half of deliveries occurring in the home. Channon describes mothers' perception of birth size as being influenced by various neighborhood and regional factors within a societal context that frames a reference for how mothers gauge their child's size.[33] Applied to our study population, children in this community relative to the global context are generally smaller perhaps leading mothers to gauge smaller children as being of average size.

We observed both low individual-level accuracy and high population-level bias for the preterm birth indicator generated from maternal reports of length of gestation at birth. Several studies have reported high degrees of accuracy of gestational age reports from mothers in developed countries.[19,21,22,34,35] One study conducted in the US Nurses' Health Studies population reported moderate sensitivity (68%) and high specificity (92%) using maternally reported gestational age to classify preterm birth.[20] Our study is limited in that we did not ask mothers to provide a numerical estimate of gestational age and have only reported the validity of using categories of gestational length at birth. We modeled this question after the format of the birth size question used in the DHS and MICS surveys, and since this question has not been used outside of this study, we recommend it be further refined before use in other settings.

Contrary to our hypothesis, the length of time since delivery did not affect the validity of maternal report for LBW and longer length of time among mothers we visited 20 months since birth and later resulted in improved accuracy for preterm birth. Some studies have reported improved accuracy and agreement between medical records and maternal report for birthweight and gestational age associated with shorter length of recall[21,34,36] while others have found accuracy of maternal report does not significantly deteriorate over time.[20,24,26] All these studies investigated patterns over longer periods of time spanning years rather than months, which limits comparability to our study findings. We observed a slightly lower degree of accuracy of maternal report used to correctly classify LBW for female compared to male children, which was not observed in other studies in developing countries.[26,27] Researchers have documented the association of sex biases in neonatal care-seeking, household food allocation and higher mortality among girls compared to boys and the persistence of son preferences in South Asia.[37–39] Further work is needed to explore whether this bias may be applicable to the accuracy of maternal report in this setting. We also observed slight improvements in maternal report accuracy associated with maternal education, consistent other studies' findings.[19,22,26,27] Finally, across all three indicators, we found that multiparous mothers had greater accuracy compared to first-time mothers, which contrasts with patterns described in prior studies.[19,26,27,35]

A strength of our study is the inclusion of home births in a population with relatively high prevalence of LBW and preterm birth. In the South Asian region, where around 69% of newborns are not weighed at birth, perhaps mothers place less importance on remembering and documenting birthweight, as evidenced by very few mothers who were able to present a birthweight card. Also, our study used accurate and calibrated scales of research quality and trained and supervised data collectors, in contrast to many delivery facilities. We also demonstrated that these indicators may be increasingly vulnerable to being underestimated in populations with higher prevalences. A limitation of our birthweight measurements used as the gold standard is that newborns were weighed up to 72 hours after birth. In this time period, newborns normally lose weight before patterns of growth and weight gain are observed. Therefore, our measurements were likely taken at a nadir and overestimate the prevalence of LBW; however, our intention was to validate maternal report rather than provide an estimate of prevalence. In addition, for home births, we are fairly confident that mothers were reporting the birthweight measured during the parent trial since this would have been the only birthweight provided to them.

However, mothers who delivered in a facility may have had their child weighed both at the facility and during participation in the parent trial. For facility births, we assumed the mother was reporting the weight measurement during the parent trial. Lastly, we did not ask mothers to report birthweight immediately after the measurement taken, which would have provided more information about whether mothers could retain birthweight information if the event occurred just prior to our interview.

Our conclusions regarding appropriate classification of preterm birth are limited since we did not ask mothers to report a numerical length of gestational age. Our categories of gestational length at birth were adapted from the DHS and MICS birth size question and have only been used in this study. Additionally, our gold standard for gestational age is based only on reported LMP, which frequently overestimates gestational age by a few days when compared to the gold standard of ultrasound measurements taken in the first trimester.[40] In low-income settings, ultrasound is generally not feasible. Our reported LMP was collected within a five-week period during pregnancy so as to optimize accurate recall, but we recognize this date is likely subject to errors in reporting.

3.6 Conclusion

The use of maternal reporting may underestimate and bias indicators for LBW and preterm birth. Additional approaches may be needed to correct for these inaccuracies in large surveys and improve methods to monitor and track progress towards global newborn health targets. The findings of this study may have limited generalizability to settings with high prevalence of LBW and preterm births and where the majority of births take place in the home. Further work is needed to explore whether these conclusions on the validity of maternal reporting hold in similar rural and low-income settings. Additional studies may be needed to understand the range of likely individual-level accuracy and population bias for these indicators in settings where women more commonly deliver in facilities and health cards are more frequently retained.

3.7 References

1. United Nations Children's Fund and World Health Organization. Low Birthweight: Country, regional and global estimates. 2004.

2. United Nations Children's Fund (UNICEF). Improving child nutrition. The achievable imperative for global progress. New York, NY: 2013.

3. Risnes KR, Vatten LJ, Baker JL, et al. Birthweight and mortality in adulthood: A systematic review and meta-analysis. Int J Epidemiol 2011;40(3):647–61.

4. WHO. Global Nutrition Targets 2025: Low Birth Weight Policy Brief. Geneva, Switzerland: 2014.

5. Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. Lancet 2016;388(10063):3027-35.

6. Howson CP, Kinney M V, McDougall L, Lawn JE. Born too soon: preterm birth matters. Reprod Health 2013;10 Suppl 1:S1.

7. March of Dimes, PMNCHS, Save the Children, WHO. Born Too Soon: The Global Action Report on Preterm Birth. Eds CP Howson, MV Kinney, JE Lawn. Geneva: 2012.

8. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: A systematic analysis and implications. Lancet2012;379(9832):2162–72.

9. Unicef. Committing to Child Survival : A Promise Renewed. Unicef Progress Report 2014. 2014.

10. Villar J, Abalos E, Carroli G, et al. Heterogeneity of perinatal outcomes in the preterm delivery syndrome. Obstet Gynecol 2004;104(1):78–87.

11. Alexander GR. "Prematurity at Birth: Determinants, Consequences, and Geographic Variation." Preterm Birth: Causes, Consequences, and Prevention. Washington D.C.: Institute of Medicine of the National Academies; 2007.

12. Hancioglu A, Arnold F. Measuring Coverage in MNCH: Tracking Progress in Health for Women and Children Using DHS and MICS Household Surveys. PLoS Med 2013;10(5).

13. Central Bureau Of Statistics/Government of Nepal. National Population and Housing Census 2011. Kathmandu, Nepal: 2012.

14. Vecchio TJ. Predictive value of a single diagnostic test in unselected populations.

N Engl J Med 1966;274(21):1171–3.

15. 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).

16. Stanton CK, 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).

17. Blanc AK, Warren C, Mccarthy KJ, Kimani J, Ndwiga C, Ramarao S. Assessing the validity of indicators of the quality of maternal and newborn health care in Kenya. J Glob Health 2016;6(1):1–13.

18. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J Epidemiol 2004;159(7):702–6.

19. Adegboye ARA, Heitmann BL. Accuracy and correlates of maternal recall of birthweight and gestational age. BJOG An Int J Obstet Gynaecol 2008;115(7):886–93.

20. Tomeo CA, Rich-Edwards JW, Michels KB, et al. Reproducibility and validity of maternal recall of pregnancy-related events. Epidemiology. 1999;10(6):774–7.

21. Sou SC, Chen WJ, Hsieh WS, Jeng SF. Severe obstetric complications and birth characteristics in preterm or term delivery were accurately recalled by mothers. J Clin Epidemiol 2006;59(4):429–35.

22. Bat-Erdene U, Metcalfe A, McDonald SW, et al. Validation of Canadian mothers' recall of events in labour and delivery with electronic health records. BMC Pregnancy Childbirth 2013;13(Suppl 1):S3.

23. O'Sullivan JJ, Pearce MS, Parker L. Parental recall of birth weight: How accurate is it? Arch Dis Child 2000;82(3):202–3.

24. Li CY, Wei JN, Lu TH, Chuang LM, Sung FC. Mothers tended to overreport categorical infant birth weight of their children. J Clin Epidemiol 2006;59(12):1319–25.

25. McCarthy KJ, Blanc AK, Warren CE, Kimani J, Mdawida B, Ndwidga C. Can surveys of women accurately track indicators of maternal and newborn care? A validity and reliability study in Kenya. J Glob Health 2016;6(2).

26. Boeke CE, Marín C, Oliveros H, Mora-Plazas M, Agudelo-Cañas S, Villamor E. Validity of maternal birthweight recall among colombian children. Matern Child Health J 2012;16(4):753–9.

27. Lule S a., Webb EL, Ndibazza J, et al. Maternal recall of birthweight and birth size

in Entebbe, Uganda. Trop Med Int Heal 2012;17(12):1465–9.

28. Araújo CL, Dutra CL, Hallal PC. Validity of maternal report on birth weight 11 years after deliver: the 1993 Pelotas Birth Cohort Study, Rio Grande do Sul State, Brazil. Cad Saude Publica 2007;23(10):2421–7.

29. Boerma JT, Weinstein KI, Rutstein SO, Sommerfelt AE. Data on birth weight in developing countries: Can surveys help? Bull World Health Organ 1996;74(2):209–16.

30. Mbuagbaw L, Gofin R. Can recall of birth size be used as a measure of birthweight in Cameroon? Paediatr Perinat Epidemiol 2010;24(4):383–9.

31. Eggleston E, Tsui AO, Fortney J. Assessing survey measures of infant birth weight and birth size in Ecuador. J Biosoc Sci 2000;32(3):373–82.

32. Islam MM. Can maternal recalled birth size be used as a proxy measure of birth weight? An evaluation based on a population health survey in Oman. Matern Child Health J 2014;18(6):1462–70.

33. Channon AAR. Can mothers judge the size of their newborn? Assessing the determinants of a mother's perception of a baby's size at birth. J Biosoc Sci 2011;43(5):555–73.

34. Hakim RB, Tielsch JM, See LC. Agreement between maternal interview- and medical record-based gestational age. Am J Epidemiol 1992;136(5):566–73.

35. Poulsen G, Kurinczuk JJ, Wolke D, et al. Accurate reporting of expected delivery date by mothers 9 months after birth. J Clin Epidemiol 2011;64(12):1444–50.

36. Seidman DS, Slater PE, Ever-Hadani P, Gale R. Accuracy of mothers' recall of birthweight and gestational age. Br J Obstet Gynaecol 1987;94(8):731–5.

37. Rosenstock S, Katz J, Mullany LC, et al. Sex differences in morbidity and care-seeking during the neonatal period in rural southern Nepal. J Heal Popul Nutr 2015;33(1):1–10.

38. Chen LC, Huq E, Souza SD. Sex bias in the family allocation of food and health care in rural Bangladesh. Popul Dev Rev 1981;7(1):55–70.

39. Gupta M Das, Jiang Z, Li B, Xie Z, Chung W, Bae H. Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea. J Dev Stud 2003;40(2):153–87.

40. Hoffman CS, Messer LC, Mendola P, Savitz DA, Herring AH, Hartmann KE. Comparison of gestational age at birth based on last menstrual period and ultrasound during the first trimester. Paediatr Perinat Epidemiol 2008;22(6):587–96.

3.8 Tables and Figures

Table 3.1: Birthweight, birth size, and gestational age birth timing questions administered

	Question	Answer choices
1.	When your child was born, was he/she born very early, early, on time, late, or very late?	1=Very early 2=Early 3=On time 4=Late 5=Very late 9=Don't know
2.	When your child was born, was he/she very large, larger than average, average, smaller than average, or very small?	1=Very small 2=Smaller than average 3=Average 4=Larger than average 5=Very large 9=Don't know
3.	Was your child weighed at birth? (Within 1 hour after birth; including a Balposan worker)	0=No (Go to 3a) 1=Yes (Go to 3b) 9=Don't Know (Go to 3a)
3a.	Was your child weighed within 3 days after birth?	0=No (Go to next question) 1=Yes (Go to 3b) 9=Don't Know (Go to next question)
3b.	How much did your child weigh the first time they were weighed? (ASK MOTHER TO RECALL ONLY; DO NOT USE BIRTH RECORD OR CERTIFICATE)	(gm)_______________ 9999=Don't Know
3c.	Do you have a birth record or certificate with your child's birth weight recorded the first time they were weighed?	0=No (Go to next question) 1=Yes (Go to 3d) 9=Don't Know (Go to next question)
3d.	(CHECK THE BIRTH RECORD OR CERTIFICATE AND RECORD THE CHILD'S BIRTHWEIGHT. RECORD WHETHER HOSPITAL OR PARENT TRIAL CARD)	(gm) _______________ 1=Hospital card 2=Parent trial card

Figure 3.1: Flowchart for participant selection

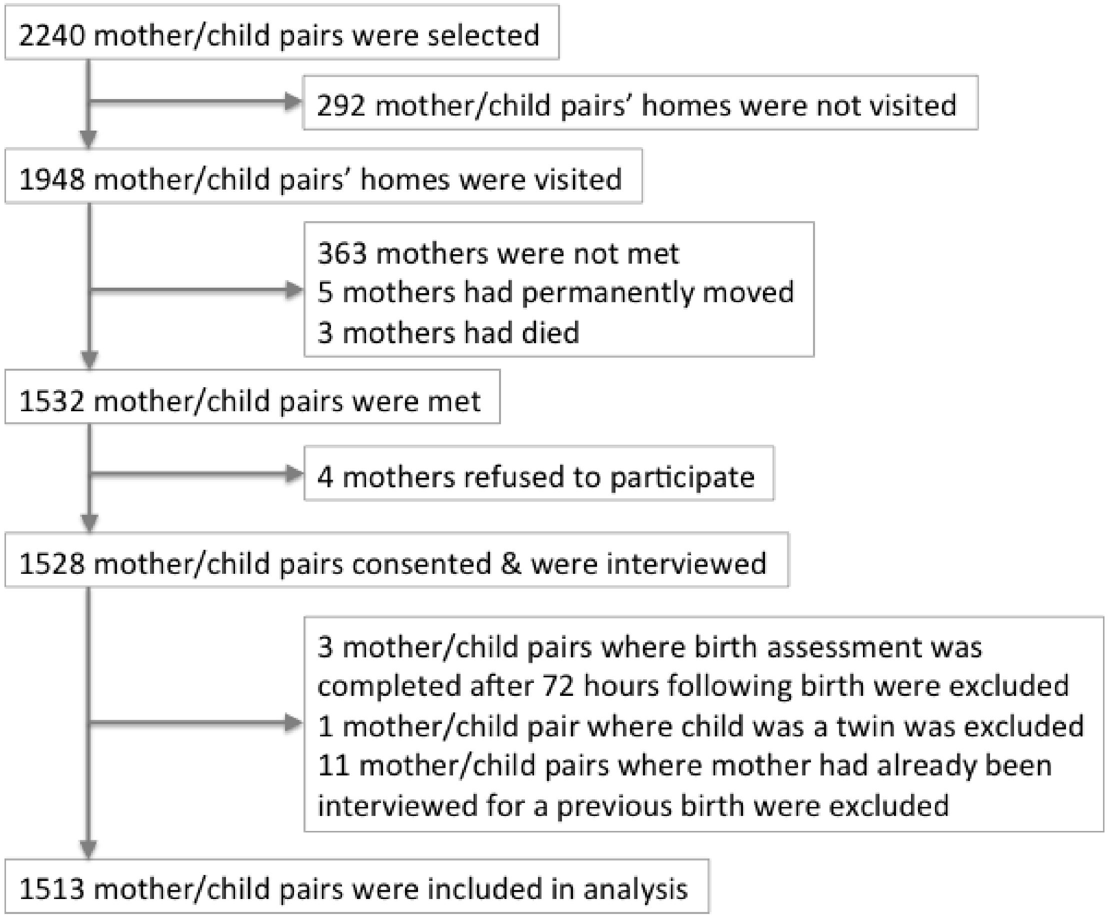

Table 3.2: Characteristics of selected mothers & children

	N	% or Mean	95% CI
Child Age (months)	1513	10.7	(10.3, 11.1)
Child Sex	1513		
Male	839	55.5	(52.9, 57.9)
Female	674	44.5	(42.1, 47.1)
Place of delivery	1512		
Home	755	53.8	(51.3, 56.3)
Facility	662	46.2	(43.7, 48.7)
Maternal Age (yrs)	1513		
<19	253	16.7	(14.9, 18.7)
20-29	1045	69.1	(66.7, 71.3)
30-39	197	13.0	(11.4, 14.8)
>40	18	1.2	(0.7, 1.9)
Maternal Education	1513		
No Schooling	1035	68.4	(65.9, 70.7)
Any Schooling	478	31.6	(29.3, 34.0)
Parity	1512		
primiparous	431	28.5	(26.3, 30.9)
second	385	25.5	(23.3, 27.7)
third	312	20.6	(18.7, 22.7)
fourth or higher	384	25.4	(23.2, 27.6)
Ethnicity	1512		
Pahadi	57	3.8	(2.9, 4.9)
Madhesi	1455	96.2	(95.1, 97.1)
HH latrine status	1512		
No latrine	1077	71.2	(68.9, 73.5)
Had latrine	435	28.8	(26.5, 31.1)
HH electricity	1512		
No electricity	298	19.7	(17.7, 21.8)
Had electricity	1214	80.3	(78.2, 82.2)
Land ownership	1512		
Did not own land	39	2.6	(1.8, 3.5)
Owns land	1473	97.4	(96.5, 98.1)

Figure 3.2: Distributions of measured and reported birthweight

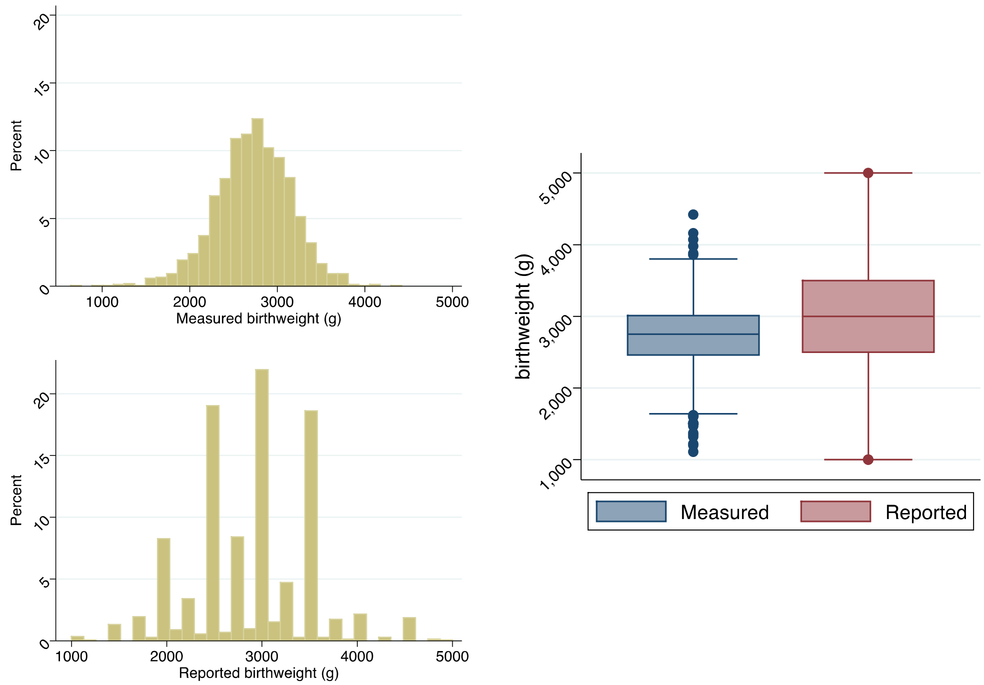

Figure 3.3: Distribution of gestational age

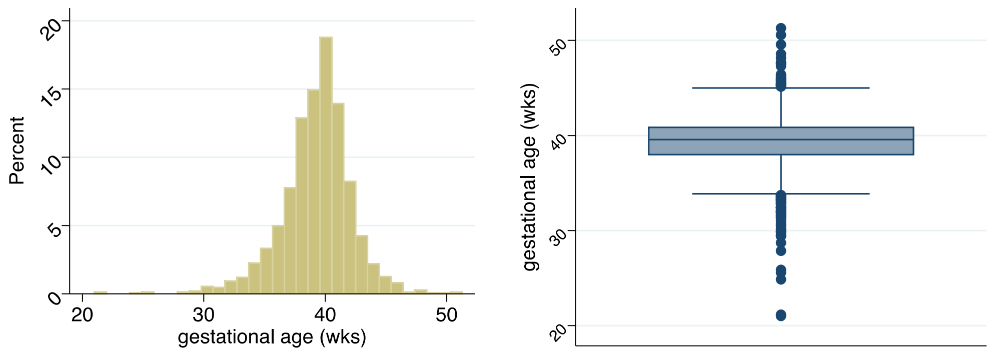

Table 3.3: Summary of birthweight and low birthweight

	Measured Birthweight			Reported Birthweight		
	N	Mean (SD) grams	LBW (< 2500g) % (95%CI)	N	Mean (SD) grams	LBW (< 2500g) % (95%CI)
All	1497	2726 (435)	27.7 (25.4, 30.0)	1439	2886 (608)	17.1 (15.2, 19.1)
Sex						
Male	831	2791 (445)	23.2 (20.5, 26.2)	795	2926 (612)	16.5 (14.1, 19.2)
Female	666	2644 (407)	33.2 (29.7, 36.9)	604	2836 (601)	17.9 (15.1, 21.0)
Birth Size						
Very small	31	1947 (612)	80.6 (62.7, 91.2)	27	2020 (827)	70.4 (50.6, 84.7)
Small	84	2298 (411)	64.3 (53.5, 73.8)	78	2227 (471)	65.4 (54.1, 75.1)
Average	1260	2748 (386)	25.0 (22.7, 27.5)	1218	2905 (553)	13.6 (11.8, 15.7)
Large	104	2985 (481)	17.3 (11.2, 25.9)	102	3298 (619)	8.8 (4.6, 16.1)
Very large	16	3078 (334)	6.3 (0.8, 35.1)	13	3577 (793)	0
Don't know	2	2430 (156)	50.0 (1.9, 98.1)	1	2000 (-)	1

Table 3.4: Summary of gestational age and preterm birth

	Gestational Age (weeks)		
	N	Mean (SD)	Preterm (<37 weeks) % (95%CI)
All	1507	39.3 (2.9)	16.1 (14.3, 17.6)
Sex			
Male	834	39.2 (3.0)	16.7 (14.3, 19.4)
Female	673	39.4 (2.8)	15.5 (12.9, 18.4)
Length of Pregnancy			
Very early	24	35.0 (5.2)	58.3 (37.9, 76.2))
Early	61	37.4 (3.1)	36.1 (25.0, 48.9)
On time	1278	39.3 (2.8)	15.6 (13.7, 17.7)
Late	115	40.9 (2.1)	3.5 (1.3, 8.9)
Very late	23	40.8 (3.6)	8.7 (2.1, 29.6)
Don't know	6	37.2 (0.9)	33.3 (7.2, 76.3)

Table 3.5: Birthweight cards by place of delivery

	Total N	Home n (%)	Facility n (%)
Mother was able to present a card with a birthweight	1486	75 (9.4)	53 (7.7)
Total cards presented	128	75	53
Card was from:			
Facility	22	0	22 (100)
Study	106	75 (70.8)	31 (29.2)
Low birth weight using measured birthweight	128	13 (17.3)	14 (26.4)
Low birth weight using birthweight on card	128	15 (20.0)	12 (22.6)

Table 3.6: Overall sensitivity, specificity, AUC, IF

Indicator	n	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	AUC (95% CI)	'True' prevalence (%) (95% CI)	Estimated survey-based prevalence (%) (95% CI)	IF (95% CI)
a) LBW using reported birthweight	1434	45.0 (40.0 - 50.1)	93.5 (91.8 - 94.9)	0.69 (0.67 - 0.72)	27.3 (25.0 – 30.0)	17.0 (15.1 – 19.1)	0.62 (0.52 - 0.72)
b) LBW using reported birth size	1497	19.1 (15.4 - 23.2)	96.7 (95.4 - 97.7)	0.58 (0.56 - 0.60)	27.7 (25.5 – 30.1)	7.7 (6.4 – 9.1)	0.28 (0.22 - 0.34)
c) PTB using reported length of gestation	1507	14.8 (10.6 - 19.9)	96.1 (94.9 - 97.1)	0.56 (0.53 - 0.58)	16.1 (14.3 – 18.1)	5.7 (4.6 – 7.0)	0.35 (0.27 - 0.44)

Table 3.7: 2x2 tables for low birthweight and preterm birth indicators

A) Low birth weight using reported birthweight				
		LBW by measured birthweight		
		Yes	No	
LBW by reported birthweight	Yes	176	68	244
	No	215	975	1190
		391	1043	1434

B) Low birth weight using reported birth size				
		LBW by measured birthweight		
		Yes	No	
LBW by reported birth size	Yes	79	36	115
	No	335	1047	1382
		414	1083	1497

C) Preterm using reported birth timing				
		Preterm by calculated gestational age		
		Yes	No	
Preterm by reported length of pregnancy	Yes	36	49	85
	No	207	1215	1422
		243	1264	1507

Figure 3.4: Inflation factors

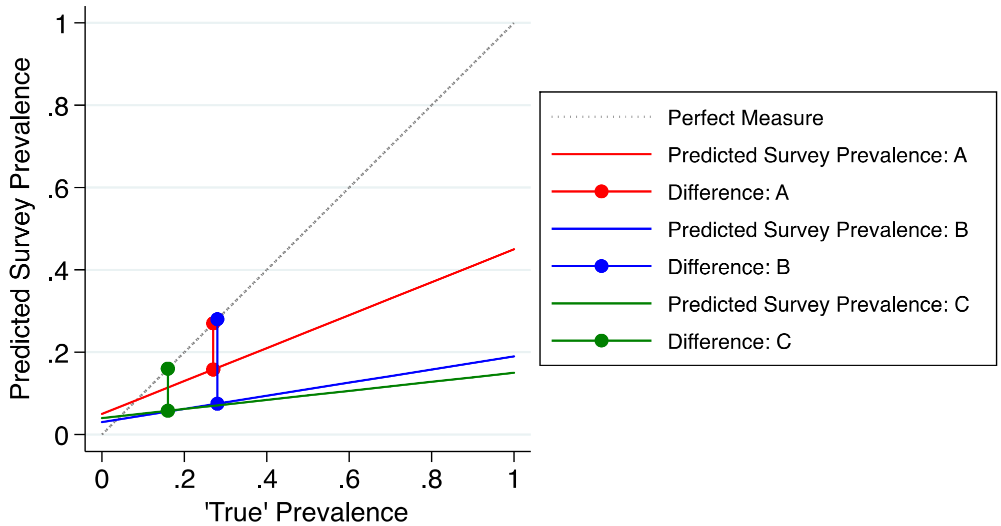

Figure 3.4: Difference between measured versus 'true' prevalence of A) LBW using reported birthweight, B) LBW using reported birth size, C) preterm birth using reported length of pregnancy

Table 3.8: Sensitivity, specificity, AUC stratified by child sex, place of delivery, maternal education, maternal age, and parity

Indicator	Total N	Sensitivity	(95% CI)	Specificity	(95% CI)	AUC	(95% CI)
Low birth weight using reported birthweight	1434	45.0%	[40.0 - 50.1]	93.5%	[91.8 - 94.9]	0.69	[0.67 - 0.72]
By child sex							
Male	793	48.9%	[41.4 - 56.4]	93.1%	[90.9 - 95.0]	0.71	[0.67 - 0.75]
Female	641	41.7%	[35.0 - 48.7]	94.0%	[91.3 - 96.0]	0.68	[0.64 - 0.71]
By place of delivery							
Home	666	38.9%	[32.2 - 46.0]	94.8%	[92.4 - 96.7]	0.67	[0.63 - 0.70]
Facility	767	51.6%	[44.2 - 58.9]	92.4%	[89.9 - 94.4]	0.72	[0.68 - 0.76]
By maternal education							
None	968	47.6%	[41.6 - 53.7]	92.2%	[90.0 - 94.1]	0.70	[0.67 - 0.73]
Any	466	38.8%	[29.9 - 48.3]	96.0%	[93.4 - 97.8]	0.67	[0.63 - 0.72]
By maternal age							
<20 yrs	240	46.9%	[36.6 - 57.3]	94.4%	[89.3 - 97.6]	0.71	[0.65 - 0.76]
20+ yrs	1194	44.4%	[38.6 - 50.3]	93.3%	[91.5 - 94.9]	0.69	[0.66 - 0.72]
By parity							
Primiparous	414	46.4%	[38.6 - 54.3]	94.4%	[90.7 - 96.9]	0.70	[0.66 - 0.74]
2 or more children	1020	44.0%	[37.4 - 50.8]	93.2%	[91.2 - 94.9]	0.69	[0.65 - 0.72]
Low birth weight using reported birth size	1497	19.1%	[15.4 - 23.2]	96.7%	[95.4 - 97.7]	0.58	[0.56 - 0.60]
By sex							
Male	831	20.2%	[14.8 - 26.6]	96.6%	[94.8 - 97.8]	0.58	[0.56 - 0.76]
Female	666	18.1%	[13.3 - 23.8]	96.9%	[94.8 - 98.3]	0.58	[0.56 - 0.60]

Indicator	Total N	Sensitivity	(95% CI)	Specificity	(95% CI)	AUC	(95% CI)
By place of delivery							
Home	691	17.8%	[12.9 - 23.7]	96.4%	[94.4 - 97.9]	0.57	[0.54 - 0.60]
Facility	805	20.4%	[15.1 - 26.6]	96.9%	[95.1 - 98.1]	0.59	[0.56 - 0.62]
By maternal education							
None	1022	21.3%	[16.8 - 26.4]	96.4%	[94.8 - 97.6]	0.59	[0.56 - 0.61]
Any	475	13.6%	[8.0 - 21.1]	97.2%	[94.9 - 98.6]	0.55	[0.52 - 0.59]
By maternal age							
<20 yrs	250	20.8%	[13.4 - 30.0]	96.0%	[91.4 - 98.5]	0.58	[0.54 - 0.63]
20+ yrs	1247	18.5%	[14.4 - 23.3]	96.8%	[95.4 - 97.8]	0.58	[0.55 - 0.60]
By parity							
Primiparous	425	17.8%	[12.3 - 24.4]	96.5%	[93.4 - 98.4]	0.57	[0.54 - 0.60]
2 or more children	1072	20.0%	[15.2 - 25.6]	96.7%	[95.3 - 97.8]	0.58	[0.56 - 0.61]
Preterm using reported birth timing	1507	14.8%	[10.6 - 19.9]	96.1%	[94.9 - 97.1]	0.56	[0.53 - 0.58]
By sex							
Male	834	14.4%	[9.0 - 21.3]	96.7%	[95.1 - 97.9]	0.56	[0.53 - 0.59]
Female	673	15.4%	[9.1 - 23.8]	95.4%	[93.4 - 97.0]	0.55	[0.52 - 0.59]
By place of delivery							
Home	695	14.7%	[8.3 - 23.5]	95.2%	[93.1 - 96.7]	0.55	[0.51 - 0.59]
Facility	811	14.9%	[9.6 - 22.6]	97.0%	[95.4 - 98.1]	0.56	[0.53 - 0.59]
By maternal education							
None	1029	12.2%	[7.9 - 17.8]	98.0%	[96.8 - 98.8]	0.55	[0.53 - 0.58]
Any	478	23.6%	[13.2 - 37.0]	92.4%	[89.5 - 94.8]	0.58	[0.52 - 0.64]
By maternal age							
<20 yrs	253	33.3%	[19.1 - 50.2]	94.4%	[90.4 - 97.1]	0.64	[0.56 - 0.72]

Indicator	Total N	Sensitivity	(95% CI)	Specificity	(95% CI)	AUC	(95% CI)
20+ yrs	1254	11.3%	[7.3 - 16.4]	96.5%	[95.2 - 97.5]	0.54	[0.52 - 0.56]
By parity							
Primiparous	432	19.7%	[10.9 - 31.3]	92.9%	[89.8 - 95.3]	0.56	[0.51 - 0.61]
2 or more children	1075	13.0%	[8.4 - 18.8]	97.4%	[96.2 - 98.4]	0.55	[0.53 - 0.58]

Figure 3.5: Sensitivity and specificity of LBW and preterm births over recall time

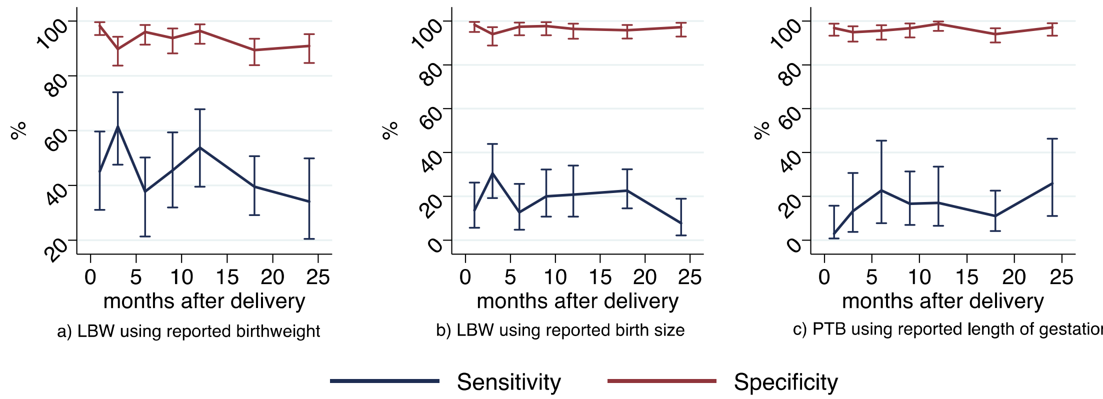

Table 3.9: Sensitivity, specificity, AUC for LBW using reported birthweight by recall group

Recall Group (months after birth)	Total N	Sensitivity	(95% CI)	Specificity	(95% CI)	AUC	(95% CI)
1 Month	220	45.1%	(31.1 - 59.7)	98.2%	(94.9 - 99.6)	0.72	(0.65 - 0.79)
3 Months	204	61.4%	(47.6 - 74.0)	89.8%	(83.7 - 94.2)	0.76	(0.69 - 0.82)
6 Months	195	34.8%	(21.4 - 50.2)	96.0%	(91.4 - 98.5)	0.65	(0.58 - 0.73)
9 Months	185	45.5%	(32.0 - 59.4)	93.8%	(88.2 - 97.3)	0.70	(0.63 - 0.77)
12 Months	189	53.8%	(39.5 - 67.8)	96.4%	(91.7 - 98.8)	0.75	(0.68 - 0.82)
18 Months	265	39.5%	(29.2 - 50.7)	89.4%	(83.9 - 93.5)	0.65	(0.59 - 0.70)
24 Months	176	34.1%	(20.5 - 49.9)	90.9%	(84.7 - 95.2)	0.63	(0.55 - 0.70)

Table 3.10: Sensitivity, specificity, AUC, IF for LBW using reported birth size by recall group

Recall Group (months after birth)	Total N	Sensitivity	(95% CI)	Specificity	(95% CI)	AUC	(95% CI)
1 Month	224	13.7%	(5.7 - 26.3)	98.3%	(95.0 - 99.6)	0.56	(0.51 - 0.61)
3 Months	208	30.5%	(19.2 - 43.9)	94.0%	(88.8 - 97.2)	0.62	(0.56 - 0.69)
6 Months	202	12.8%	(4.8 - 25.7)	97.4%	(93.5 - 99.3)	0.55	(0.50 - 0.60)
9 Months	193	20.0%	(10.8 - 32.3)	97.7%	(93.5 - 99.5)	0.59	(0.54 - 0.64)
12 Months	193	20.8%	(10.8 - 34.1)	96.4%	(91.9 - 98.8)	0.59	(0.53 - 0.64)
18 Months	285	22.6%	(14.6 - 32.4)	95.8%	(92.0 - 98.2)	0.59	(0.55 - 0.64)
24 Months	192	7.8%	(2.2 - 18.9)	97.2%	(92.9 - 99.2)	0.53	(0.49 - 0.57)

Table 3.11: Sensitivity, specificity, AUC, IF for preterm using reported length of pregnancy by recall group

Recall Group (months after birth)	Total N	Sensitivity	(95% CI)	Specificity	(95% CI)	AUC	(95% CI)
1 Month	223	3.0%	(0.8 - 15.8)	96.8%	(93.3 - 98.8)	0.50	(0.47 - 0.53)
3 Months	207	13.3%	(3.8 - 30.7)	94.9%	(90.6 - 97.6)	0.54	(0.48 - 0.61)
6 Months	203	22.7%	(7.8 - 45.4)	95.6%	(91.5 - 98.1)	0.59	(0.50 - 0.68)
9 Months	195	16.7%	(7.0 - 31.4)	96.7%	(92.5 - 98.9)	0.57	(0.51 - 0.63)
12 Months	192	17.1%	(6.6 - 33.6)	98.7%	(95.5 - 99.8)	0.58	(0.52 - 0.64)
18 Months	288	11.1%	(4.2 - 22.6)	94.0%	(90.2 - 96.7)	0.53	(0.48 - 0.57)
24 Months	199	25.9%	(11.1 - 46.3)	97.1%	(93.3 - 99.0)	0.62	(0.53 - 0.70)

Table 3.12: Correct classification of newborns as low birthweight

	A				B			
	n	aRR	95% CI	p-value	n	aRR	95% CI	p-value
Time since birth/Child age (months)	1433	0.99	(0.99, 1.00)	0.003	1495	0.99	(0.99, 1.00)	0.09
Child sex								
Male (ref)	793				830			
Female	640	0.93	(0.88, 0.98)	0.006	665	0.90	(0.84, 0.95)	<0.001
Place of delivery								
Home (ref)	767				804			
Facility	666	0.95	(0.90, 1.01)	0.09	691	0.96	(0.90, 1.02)	0.16
Maternal age (yrs)	1433	1.00	(0.99, 1.01)	0.44	1495	1.00	(1.00, 1.01)	0.35
Maternal education								
None (ref)	968				1021			
Any	465	1.06	(1.00, 1.11)	0.05	474	1.09	(1.02, 1.16)	0.02
Parity								
Primiparous (ref)	413				424			
Second child	371	1.09	(1.01, 1.18)	0.03	382	1.21	(1.10, 1.32)	<0.001
Third child	292	1.09	(1.00, 1.20)	0.05	311	1.23	(1.11, 1.36)	<0.001
Fourth child or greater	357	1.07	(0.96, 1.19)	0.23	378	1.21	(1.07, 1.36)	0.002
Ethnicity								
Pahadi (ref)	57				57			
Madhesi	1376	1.02	(0.89, 1.17)	0.80	1438	0.96	(0.83, 1.11)	0.59

Table 3.12: Modified Poisson regression with robust error variance to estimate risk ratio of correctly classifying newborns as A) LBW using reported birthweight and B) LBW using reported birth size, adjusting for child sex, place of delivery, maternal age, maternal education, parity, and ethnicity

Table 3.13: Correct classification of newborns as preterm birth

		C		
	n	**aRR**	**95% CI**	**p-value**
Time since birth/Child age (1-20 months)	1505	1.00	(0.99, 1.00)	0.12
Time since birth/Child age (>20 months)	1505	1.02	(1.01, 1.04)	0.001
Child sex				
Male (ref)	833			
Female	672	1.00	(0.96, 1.05)	0.93
Place of delivery				
Home (ref)	810			
Facility	695	1.03	(0.98, 1.08)	0.19
Maternal age (yrs)	1505	0.99	(0.98, 1.00)	0.003
Maternal education				
None (ref)	1028			
Any	477	1.01	(0.96, 1.07)	0.64
Parity				
Primiparous (ref)	431			
Second child	383	1.08	(1.02, 1.16)	0.02
Third child	311	1.14	(1.06, 1.23)	0.001
Fourth child or greater	380	1.11	(1.01, 1.22)	0.03
Ethnicity				
Pahadi (ref)	57			
Madhesi	1448	0.88	(0.81, 0.95)	0.002

Table 3.13: Modified Poisson regression with robust error variance to estimate risk ratio of correctly classifying newborns as C) preterm term using relative length of pregnancy, adjusting for child sex, place of delivery, maternal age, maternal education, parity, and ethnicity

Chapter 4: Mothers' perception of questions assessing birth size and length of pregnancy in rural Nepal

Karen T. Chang[1], Luke C. Mullany[1], Subarna K. Khatry[2], Steven C. LeClerq[1,2], Melinda K. Munos[1], Joanne Katz[1]

[1]Johns Hopkins University Bloomberg School of Public Health, Department of International Health, 601 N. Wolfe Street, Baltimore, MD, 21205, USA; [2]the Nepal Nutrition Intervention Project–Sarlahi, Lalitpur, Nepal

4.1 Abstract

Background

Quantitative validation studies alone may not be able to distinguish between instances when participants did not accurately report an event versus when participants did not understand a question. We used a qualitative study design to acquire an in-depth understanding about how mothers perceive the phrasing of questions assessing birth size and length of gestation.

Methods

We conducted two focus group discussions (FGDs) with study staff who administered a quantitative questionnaire and 12 in-depth interviews (IDIs) with mothers who had participated in this quantitative component. Transcripts were coded and analyzed for themes in patterns of meaning within and across FGDs and IDIs. Using this thematic map, we synthesized our data into common and divergent responses from participants to facilitate our interpretation of the quantitative findings.

Findings

We identified five themes specific to this analysis: difficulties with the length of pregnancy question, challenges in administering the birth size question, the perceived effect of time since birth on mothers' ability to remember information, the language and style differences specific to this setting, and the study context shaping the relationship between study staff and mothers who participated and how this may have influenced mothers' responses. Visual aids may help to scale the question about birth size within a cultural frame of reference for maternal reports to be more interpretable. Among both study staff and mothers, a longer period of time since the birth of a child was thought to be associated with diminished accuracy of maternal reports, a perception not supported by our quantitative findings.

Interpretation

Poor validity of low birth weight (LBW) and preterm birth indicators based on maternal reports may be partly attributed to challenges in maternal understanding of questions assessing birth size and length of pregnancy. Additional research is needed to confirm these findings regarding maternal comprehension and to further evaluate the utility of visual aids developed for this study.

4.2 Introduction

Low birthweight (LBW, <2500g) and preterm birth (<37 weeks) are associated with increased risk of child mortality, severe disability, cognitive impairment, and other long-term health problems.[1–4] Worldwide, about 20 million LBW infants are born annually, and in South Asia, a quarter of all live births are LBW.[2,5] Each year, approximately 15 million preterm newborns are born globally.[1,3,6,7] Preterm birth, disproportionately burdening South Asian and African countries, is the leading cause of neonatal deaths and the second leading cause of under-five mortality.[6,7] As part of the Sustainable Development Goals, a target of reducing neonatal and child mortality to 12 and 25 deaths per 1,000 live births, respectively, was set for all countries by 2030 Monitoring LBW and preterm birth indicators over time informs global progress towards achieving these newborn and child health targets.[8]

Maternally-reported information collected as part of national household surveys, like the Demographic and Health Surveys (DHS) and the Multiple Indicator Cluster Surveys (MICS), are often the only source of child health data available to generate birth indicators in low-income countries.[9] Under these approaches, mothers are asked to recall events related to their child's birth that might have taken place up to five years prior to administration of the survey.[9] Given our reliance on data from such surveys, efforts to evaluate the validity of maternal recall of newborn health are necessary. Quantitative validation studies, however, may not be able to distinguish between instances when participants did not accurately report an event versus when participants did not understand a question.[10–12] Feedback from data collectors themselves in conducting quantitative surveys can be useful to identify any additional probes that were provided to participants who had difficulty understanding a question and to gauge the level of consistency in administering questionnaires across data collectors.[13,14] Prior studies have investigated respondents' comprehension of survey questions that are similar to those used in DHS and MICS.[10–12] Results of these studies may help to identify questions that may be difficult for mothers to understand and methods that could improve the quality of data collected in surveys. We used a qualitative study design to acquire an in-depth understanding about how mothers perceive the phrasing of questions assessing birth size and length of gestation. We also describe the experience of study staff in administering these questions, the perceived effect of time since birth on mothers' ability to remember information, the language and style differences specific to this setting, and the study context shaping the relationship between study staff and mothers who participated and how this may have influenced our findings.

4.3 Methods

Study setting

We conducted the study in the rural Sarlahi District of Nepal, where only about half of its predominantly Hindu residents are able to read and write.[15] Over a third of residents are younger than 15 years of age, and almost one in five married women were younger than 15 years old at their first marriage.[15] The study was nested within a community-randomized trial that aimed to assess the impact of using sunflower seed oil in full-body newborn massage on neonatal morbidity and mortality (registered at ClinicalTrials.gov (NCT01177111)).

Substudy

We selected mother/child pairs from the parent trial for one additional follow-up visit using a quantitative form to ask mothers to report on circumstances of labor and delivery, immediate newborn care, postnatal care, and neonatal morbidity and care seeking in the first 7 days of life. We aimed to interview roughly the same number of mothers at each of seven follow-up times after birth: 1, 3, 6, 9, 12, 18, or 24 months. Study staff requested participation in the homes of selected mothers, administered oral consent in Nepali or Maithili and obtained a signature or thumbprint for those who agreed to participate. Specific to this quantitative analysis, mothers were asked about birthweight, birth size, length of pregnancy, and whether they had a written record of their child's birthweight. We compared maternal reports in the substudy to prospectively collected data in the parent trial (our "gold standard" estimate) to assess the validity of a) birthweight and b) birth size in correctly categorizing newborns as LBW, and c) length of pregnancy in identifying preterm births. LBW was defined as a birthweight of less than 2500 grams (gold standard), regardless of gestational age, and a birth size of "small" or "very small" (from maternal reports). Preterm birth was defined as gestational age less than 37 weeks and reported delivery of "early" or "very early." Quantitative results from this validation of maternal reports are published separately (submitted in the same supplement to the journal – fill in citation when available).

Qualitative Approach

We identified questions of interest from the quantitative form based on feedback from study staff during data collection supervision visits and on preliminary analyses of quantitative data that frequently generated discordant results comparing maternal report in the substudy to data collected in the parent trial. We focused on these questions in this qualitative analysis in order to complement and give further context to the interpretation of our quantitative findings. The questions from our quantitative interview that we further explore in this specific qualitative analysis are listed in **Table 4.1**. Logistical circumstances precluded data collection to continue until saturation; rather, we planned to complete a predefined number of FGD and IDIs prior to the start of data collection that would be logistically feasible.

From August to November 2016, we conducted focus group discussions (FGD) with the study staff who had administered the quantitative form. A discussion guide was created to cover the following themes: reflection on experiences with administering the quantitative form, identification of questions mothers had difficulty answering, discussion of reasons difficulties were encountered, description of probes used for clarification, and suggestions for how questions could be improved for better understanding. Locally resident, female qualitative interviewers were fluent in Nepali and Maithili, from the same community as our study staff, and in non-supervisory roles in an attempt to allow study staff to more openly share their experiences working on the study. We conducted the FGDs in a private room at one of the field offices during the last month of data collection for the quantitative form.

Based on information from the FGDs, we narrowed our final list of quantitative questions of interest and identified and developed visual aids for use during in-depth interviews (IDI). We restricted IDI participation in this qualitative component to 9 Village Development Committees closest to our Sarlahi study office for logistical convenience. Mothers residing in these areas who responded discordantly to at least three of the questions of interest comparing data from the parent trial to their responses in the quantitative component of the substudy were eligible to participate in IDIs. Qualitative interviewers administered oral consent in Nepali or Maithili and obtained a signature or thumbprint for mothers who agreed to participate. IDIs were conducted in a private area in households of participants. An interview guide was created to cover the following topics: willingness to discuss labor and delivery and newborn health, attitudes about newborn health checks, views about whether time since birth affects mothers' ability to remember what happened, and reflection on questions that generated discordant responses and methods to improve the accuracy of maternal responses. Figure 1 shows the dolls and photos of newborns of different sizes (A-2.2kg, B-2.6kg, C-3.1kg) developed based on suggestions from study staff during FGDs and used in administering questions related to birth size in IDIs. The FGD and IDI guides are included in **Appendices 1 and 2**. This analysis focused on questions related to birth size and length of pregnancy and the effect of time since birth on mothers' memory. Results related to views about discussing labor and delivery and newborn health with others, attitudes about newborn health checks, and other questions that generated discordant responses will be reported elsewhere.

Discussion and interview guides were created in English and translated into Nepali and Maithili by local staff. Debrief sessions were conducted with qualitative interviewers following FGDs and IDIs to reflect on the quality of the discussion/interview, summarize content, edit questions for understanding, and discuss challenges. FGD and IDIs were audio recorded and transcribed from Maithili to Nepali by the interviewers. The Nepali transcripts were then sent to translators in Kathmandu for translation to English. For any additional clarifications that were needed in the English versions of transcripts, a native Nepali speaker reviewed Nepali versions and re-translated sections as needed. Recordings, transcripts, and translations were all de-identified.

Ethical approval

The parent trial and the substudy both received ethical approval from the Johns Hopkins Bloomberg School of Public Health Institutional Review Board, Baltimore, MD, USA. Local approval was received from the Tribhuvan University Institute of Medicine, Kathmandu, Nepal for the parent trial and from the Nepal Health Research Council, Kathmandu, Nepal for the substudy.

Data analysis

We analyzed transcripts using Atlas.ti Scientific Software. Our hypotheses following preliminary analyses of quantitative data from the substudy focusing on birth size and length of pregnancy and the interview guides deductively informed the development of codes in an initial codebook. In a first round of coding, we applied both initial codes and inductively added new codes based on additional themes that arose to expand our preliminary codebook. In a second stage of the coding process, all codes were grouped into overarching axes and refined to create a final codebook applied in a second review of transcripts. We adopted a constructivist perspective in a thematic analysis to search for patterns of meaning within and across FGDs and IDIs.[16,17] Using this thematic map, we synthesized our data into common and divergent responses from participants to facilitate our interpretation of the findings.[16,17]

4.4 Results

We first conducted two FGDs, each with 6 study staff, who administered the quantitative form. All study staff were female, had at least a high school diploma and ranged from 20 to 50 years of age. In the first FGD, four of our study staff were of the Madhesi ethnicity and two were Pahadi. In the second FGD, two were Madhesi and four were Pahadi. **Table 4.2** describes characteristics of the 12 mothers who were selected for IDIs, conducted after completion of FGDs. Five themes emerged in the analysis within and across FGDs and IDIs: difficulties with the length of pregnancy question, difficulties with the birth size question, the effect of time since birth on mothers' memory, the language and style differences specific to this setting, and the relationship between staff and mothers in the quantitative interviews.

Difficulties with the length of pregnancy question

Study staff reported encountering difficulties in maternal understanding of the length of pregnancy question during both FGDs. Citing that the question is too long and the phrasing of the question is confusing, staff in an FGD also described mothers perhaps requiring more context to understand what the word, 'time,' refers to as the word was used repeatedly in translations of the question. From staff experiences, mothers occasionally misunderstood the question as asking about the time of day the child was born or the length of time they were in labor, as described below:

> Participant 2: When we ask about the time, they think we mean morning, afternoon, or night.
> Participant 3: That's what happened when we asked this question. (Laughing)
> Participant 2: Yeah, some say it happened in the morning, others at night.
> Participant 4: That's what they immediately understand by time.
> Participant 2: They specify that evening is a time, too, and that it didn't happen at night.
> Participant 3: Like Participant 1 said (pointing at Participant 1) actually, women remember the bits starting from their labor pain, and their attention is stuck there.
> Moderator: Hmm…
> Participant 3: And that's probably the reason why it's difficult for the mothers to answer the question- (Participant 2 interrupting)
> Participant 2: When we went for training, we were instructed what the phrase 'before time' means in the question. But usually, the mothers don't know what the phrase 'before time' means. They don't know what 'time' is referring to. They don't know which 'time.' Maybe that's why they get confused.

When asked about how the question could be rephrased for better comprehension and what additional probes study staff would use to help mothers understand, FGD participants suggested specifying 'preterm' or 'due date' in the following discussion:

> Participant 3: When the child was born, was he born at term or preterm… I don't think it is right. (Laughing)
> Participant 4: But this is how we explained in the field. We ask them if the child was born at term or prematurely and explain to take the difference between months and days.
> Participant 6: If the baby was born prematurely- (Participant 1 interrupted)
> Participant 1: Was your baby born after the time or before it?
> Participant 6: We asked like that.
> Participant 4: They don't understand like that. (Speaking to Participant 1) Was the baby born right at the due date or before it? Or was the baby born way before the due date or past the due date? Or way past the due date or right on time? That's how we asked and they understand it easily.

Data from IDIs with mothers supported these findings. When asked about the length of pregnancy using the original phrasing of the question, one mother described the time of day her child was born and makes reference to the possibility of requiring a caesarean-section had her child not been born within a specified time since labor:

> Participant: The doctor said that if she delivers the baby at 8 o'clock, then it is okay. Otherwise, she will give birth after a big operation. The doctor left to go home to run errands after saying that. And when she came back after running errands, the baby was delivered at 7 o'clock. The baby was not born at 8 o'clock; she was delivered at 7 o'clock. (21 year old mother)

When asking this same mother the question specifying 'preterm' as an example of 'early and specifying 'overdue' as an example of 'late,' the mother responded, her "baby was born at nine months."

Difficulties with the birth size question

Study staff were instructed to ask the birth size question exactly as phrased in the Nepali DHS and MICS surveys. Some difficulties encountered in administering this question included unfamiliarity with the local word for 'average' ('ausat') and using the local word for 'normal' ('samanya') size in clarifying probes.

> Participant 3: Maybe because they don't understand the word 'ausat' that they don't get the rest. Instead of the word 'ausat,' when we asked them how was the baby, big, small, average, in their way, they understood it.
> Participant 2: Yes. When we said smaller or bigger than average, they'd understand immediately. When asked whether [the baby] was bigger or smaller than was supposed to be, they said that it was average ['samanya'- the Nepali word for normal]. The answer comes that way…
> Moderator: Yeah… In Participant 2's (pointing at Participant 2) opinion, they wonder what the word means and how they should respond. Similarly, are there any other- (Participant 4 interrupts)
> Participant 4: When we say 'ausat' means average, they don't understand but when we just say 'average' ['samanya'- the Nepali word for normal] they understand.

In another FGD, staff explained that the question was sometimes misunderstood as asking about the height or length of the child, rather than the weight.

> Participant 6: When we ask how big was the baby then the first thing that comes to their mind is the height of the child and not the weight and therefore it becomes necessary to repeat the question indicating that it is related to weight.

When asked how mothers' understanding of this question could be improved, study staff suggested using colored photos of newborns of different sizes to show to mothers for a frame of reference.

> Participant 6: One picture/photo will not do anything. (P6 and P3 agreeing to the statement). Rather a picture of a fat child and a picture of a thin child might have been helpful. It might be helpful if the pictures were in color.
> Moderator: So you think that the photo of a fat or thin child will be helpful in understanding this question?
> (Participant 3 and Participant 6 speaking together) The photo of newly born children in color will be better for understanding of the mothers.

Based on this suggestion from the study staff, we created the visual aids in **Figure 4.1** to use in IDIs with mothers. When our qualitative interviewers asked mothers about their child's size at birth, a common response was, "My child was neither too big nor too small; he was normal." As a follow-up question, when asking a mother to then select a photo of a newborn whose size most closely resembled that of her child when he was born, one mother "looked at all the photos and then took one of the photos in her hand that was the smallest in size (Photo A) and said that her child was like that photo but was thinner than the photo." In further discussion, the mother again described her child as being of average size, "As I told you, my child was neither very big nor very small. My child was somewhere in between." This child, who was born at home, weighed 2.25kg two and a half days after birth.

Effect of time since birth on mothers' memory

Some of the study staff thought the length of time since birth did affect the accuracy of mothers' responses in some cases.

> Participant 6: The women who have a five- or six-month old child would remember. But at the beginning of data collection, there were women who had thirteen-month old children, and it was hard for them. They had even forgotten the answers to some of the questions. Others were all right.

Other staff pointed out that maternal accuracy depended more on the individual ability of the mother to remember things.

> Participant 3: So some of the women remember things even after two years while some others do not remember the things that happen within a month or so.
> Participant 2: Yes, they do not remember.
> Participant 6: This is problem of some women. All the women do not have the same memory power.
> Participant 5: All people do not have same type of brain.

This question was also posed to mothers during IDIs, and many mothers noted that their memory of events may fade with time. One mother said that day-to-day obligations and worries prevent her from remembering events at birth.

> Participant: If the mothers are free from other things and keep on thinking the same thing again and again then they can remember it. Therefore if the mothers have time to think on the time of their delivery and the baby conditions at that time then there is every chance that they can remember for longer time. But they have to engage themselves in so many other domestic chores like how to get baby educated, how to earn both times meal for family, looking after the animals and small children etc. So, most of the time their mind is occupied with these things. Don't you think that these are more important to spend time on rather than just thinking over and over about their delivery time? (19 year old mother)

Language and style differences

A frequent theme that emerged from our FGDs with study staff was the distinction between the Nepali and Maithili language and styles. In discussions related to the length of pregnancy question, one staff member referred to using the Maithili language to aid mothers' understanding, saying "Only some [mothers] won't [understand]. The mothers will understand if we explain in their language.” Another staff member in a different FGD said, “As long as we explained the question in their style (Maithili), they understood it at once, and we didn’t have to probe a lot.” In discussions related to misunderstanding of the birth size question, one of our staff explained that in “Maithili society,” mothers frequently thought the question was asking about height or length rather than weight.

Relationship between staff and mothers

In both FGDs, there was a perception that literacy and education levels of mothers were linked to the ability to understand questions. In reference to the birth size question, one of the staff explained, “When we meet literate women, when we say bigger than average and smaller than average, they understand it right away. But when we meet others, they don’t know what it means.” Later in this same discussion, another staff member shared her experience during household visits, “In fact when we go [to their houses], mothers are a bit intimidated and feel shy to talk to us.” While providing suggestions for props and pictures to use as visual aids with mothers during IDIs, study staff made a distinction between mothers who are ‘smart’ versus ‘silent.’

> Moderator: Let’s say (showing a doll) this one here. If you question the mothers showing this doll… If you use this doll to question the mothers, what do you think will happen?
> Participant 3: You have to show this to silent mothers and question.
> Participant 4: (Using the doll) This is how the navel was examined for any signs of danger in the body of the child- (P2 interrupting)
> Participant 2: If that’s the case, they won’t understand our questions at all. And if she’s silent there will be no interview. (Laughing and P4 joins)
> Moderator: Yeah…
> Participant 2: You have to show it to a mother who is smart rather than a mother who is silent.

The theme of being educated versus uneducated was also reflected in IDIs in responses from mothers related to the effect of time since birth on maternal recall.

> Participant: Sometimes the things are remembered.
> Interviewer: They are remembered?
> Participant: No, it is not all remembered. It is not written down like how the educated people do it. Nobody can remember everything. (21 year old mother)

When asked to explain how one mother understood the question about length of pregnancy, she pointed out the difference in literacy and education between herself and the qualitative interviewer.

> Interviewer: How do you understand this question? What do you think this question is trying to ask?
> Participant: What you asked me is… See, you are educated and I am illiterate. Despite that I have to use whatever wits I can gather to work. Say, I have to think about what is good and what is bad. I have to find a good path. (24 year old mother)

A final consideration that could have influenced the responses mothers gave to both our staff who administered the quantitative form and our qualitative staff is the frequent expectation mothers had that they would be compensated in some way for participating in an interview.

> Participant 6: In the area where we work, when we go there to have discussion and explain them about our objective of discussion then the only thing they ask is what will they get by participating in the discussion.
> Moderator Ok
> Participant 6: They wished to get something after participating in the discussion and therefore they used to ask this time and again during the discussion.
> Moderator: Were they asking this question after finishing your discussion?
> Participant 2: Yes
> Participant 3: Yes, they ask that
> Participant 6: Yes, they used to ask us after finishing our work because they expected to get something at the end.
> Moderator: They were expecting?
> Participant 6: Yes
> Moderator: That is fine, anything else they used to ask?
> Participant 3: In many cases, we found that they did not have a cooperative nature for our work. In some households, they used to ask this question in such a way

> that it was hard for us to answer them.
> Moderator: Is that- (Participant 3 interrupts)
> Participant 3: Yes, in some houses we felt that they were not much interested in answering our questions. When we ask questions, they would keep silent without responding to our question. So we had to ask many times, then only they answer briefly. This was not at all the places, but in some houses.

Other staff also noted that some mothers are happy to participate even if they are not compensated, saying "In some case, they seem happy in spite of not receiving anything in return for participating. They say you come and ask us how was the bath given [to the child], how did the birth take place and things like that. So, I felt they were happy although they were not getting anything." Later in this same discussion, one staff member explains that this expectation may be related to compensation earthquake victims were receiving; she said, "These days people, the earthquake victims, are getting relief fund/materials and they think that they will also get something after answering our questions."

4.5 Discussion

We previously reported low accuracy in maternal reports used to calculate LBW and preterm birth indicators as compared to birthweight and gestational age data collected as part of a community randomized trial in rural Nepal. Based on qualitative results from FGDs with study staff and IDIs with mothers who participated in the quantitative component of this study, low accuracy of maternal reports may be partly attributed to inconsistent understanding of questions related to birth size and length of pregnancy among mothers in rural Nepal. Although we had translated our quantitative forms to both

Nepali and Maithili and the local study team reviewed translations, our study identified challenges in the phrases used, and style and the length of questions administered. While we had adapted the Nepali version of the birth size question verbatim from the DHS and MICS surveys, we had created a Maithili version of the question for use in this study. The length of pregnancy question was developed using a similar sentence structure as the birth size question in English. In a study assessing the comprehension of questions in a Tanzania AIDS Indicator Survey, Yoder and Nyblade describe difficulties encountered with translation from English to Kiswahili, including problems with style and structure.[12] The authors encouraged the use of translations that are not literal, but rather, reflect the original intent of the question. Cognitive interviews may be a useful tool to gauge participant comprehension following translation of a survey from English into a local language.[18–20] Other approaches that ask different types of questions and use simplified sentence structures to measure the same construct across different cultures may also be necessary.[21] Creating a template of questions in English and simply translating them into other languages may fall short in guaranteeing equivalence in what is measured across study settings and may not be sufficient in ensuring data quality.[22]

To aid maternal understanding of the birth size question, we asked mothers during IDIs to refer to photos of newborns of varying weights and identify one that most resembled the size of her child at birth. We observed that while mothers frequently described their child as being of ‘average’ or ‘normal’ size at birth without the visual aid, mothers often selected the photo of the smallest child. Channon describes the influence of various community and regional factors within a societal context on mothers’ perception of birth size that shape a point of reference for how they assess their child's size.[23] Relative to the global context, newborns in this rural Nepali setting are generally smaller, perhaps influencing mothers to perceive smaller children as being of average size. Visual aids may assist in scaling the question about birth size within a cultural frame of reference for maternal reports to be more interpretable. Interestingly, both study staff and mothers believed accuracy of maternal report would diminish over time, consistent with our initial hypothesis; however, our quantitative findings do not support this theory in the context of generally poor maternal accuracy, even at one month after birth. Previous research has found greater accuracy and agreement between maternal reports of birthweight and gestational age and medical records associated with shorter periods of recall[24–26] while others have observed no difference over time.[27–29] However, comparability to our study is limited since these studies assessed reports following longer periods of time over years rather than months.

Adopting a constructivist paradigm, we reflect on the relationships between the study staff and mothers who participated in interviews.[30] We observed a consistent thread in both our FGDs and IDIs that suggested a power dynamic existed between the interviewer and the participant that likely influenced the types of responses collected in our study. Overall, study staff were viewed as being educated and literate while mothers who had difficulty answering questions were considered uneducated and illiterate. This dynamic may have precluded mothers from being more open in sharing their opinions because they felt intimidated or shy. It is likely that this same dynamic is operating in DHS and MICS in many countries where literacy in rural areas is low. In addition, if mothers were frequently expecting compensation for their participation, they may have provided biased responses based on what they thought interviewers wanted to hear, or they may have been less interested in engaging in in-depth conversation once they learned no compensation would be provided. As part of the parent trial, households were provided small gift items, like a baby blanket or hat, but monetary compensation was not provided. The DHS and MICS Nepal surveys also do not provide compensation for participating in interviews, so we might expect this same possible response bias.

Finally, there were several limitations in this study. We were unable to complete all 15 planned IDIs due to logistical constraints. Since we were unable to continue data collection until saturation was reached, there may also be other contributing factors for the discordant maternal responses. Transcripts were subjected to several layers of translation. Qualitative interviewers listened to audio recordings of the interviews that were primarily conducted in Maithili and directly translated these into Nepali, which may have resulted in a loss of emic terms. Although clarifications were sought from a native Nepali speaker during analysis of English transcripts, the author of this thesis is a non-native Nepali speaker, further limiting our findings from this analysis.

4.6 Conclusion

Poor validity of LBW and preterm birth indicators based on maternal response may be partly attributed to challenges in maternal understanding of questions assessing birth size and length of gestation. Findings from this qualitative study suggest specific terms in

Maithili translation and sentence structure affected maternal comprehension. Visual aids, like pictures of newborns of varying sizes, may help to scale maternal perception of birth size in specific settings. In addition, relationships and dynamics between interviewers and participants may affect the nature of responses. More work is required to further explore maternal comprehension of these questions in similar rural and low-income settings as a prelude to improving content and context in DHS and MICS surveys.

4.7 References

1. Unicef. Committing to Child Survival : A Promise Renewed. Unicef Progress Report 2014. 2014.

2. United Nations Children's Fund and World Health Organization. Low Birthweight: Country, regional and global estimates. 2004.

3. Howson CP, Kinney M V, McDougall L, Lawn JE. Born too soon: preterm birth matters. Reprod Health 2013;10 Suppl 1:S1.

4. Alexander GR. "Prematurity at Birth: Determinants, Consequences, and Geographic Variation." Preterm Birth: Causes, Consequences, and Prevention. Washington D.C.: Institute of Medicine of the National Academies; 2007.

5. United Nations Children's Fund (UNICEF). Improving child nutrition. The achievable imperative for global progress. New York, NY: 2013.

6. March of Dimes, PMNCHS, Save the Children, WHO. Born Too Soon: The Global Action Report on Preterm Birth. Eds CP Howson, MV Kinney, JE Lawn. Geneva: 2012.

7. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: A systematic analysis and implications. Lancet 2012;379(9832):2162–72.

8. Indicator and monitoring framework for the global strategy for women's, children's and adolescents' health (2016-2030). 2015.

9. Hancioglu A, Arnold F. Measuring Coverage in MNCH: Tracking Progress in Health for Women and Children Using DHS and MICS Household Surveys. PLoS Med 2013;10(5).

10. Salasibew MM, Filteau S, Marchant T. Measurement of breastfeeding initiation: Ethiopian mothers' perception about survey questions assessing early initiation of breastfeeding. Int Breastfeed J 2014;9:13.

11. Yoder PS, Rosato M, Mahmud R, et al. Women's recall of delivery and neonatal care in Bangladesh and Malawi: A study of terms, concepts and survey questions. Calverton, Maryland, USA: 2010.

12. Yoder PS, Nyblade L. Comprehension of questions in the Tanzania AIDS Indicator Survey. DHS Qual Res Stud No.10 2004.

13. Salasibew MM, Dinsa G, Berhanu D, Filteau S, Marchant T. Measurement of delayed bathing and early initiation of breastfeeding: A cross-sectional survey exploring experiences of data collectors in Ethiopia. BMC Pediatr 2015;15(1).

14. Thapa S, Dhital M, Neupane S. Assessing the quality of survey data on adolescent sexuality by talking with the field staff. Contrib Nepalese Stud 2002;29(1):1–30.

15. Central Bureau Of Statistics/Government of Nepal. National Population and Housing Census 2011. Kathmandu, Nepal: 2012.

16. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77–101.

17. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006;5(1):80–92.

18. Beatty P, Willis G. Research Synthesis: The Practice of Cognitive Interviewing. Public Opin Q 2007;71(2):287–311.

19. Giri J, Pokharel PR, Gyawali R, Bhattarai B. Translation and Validation of Modified Dental Anxiety Scale: The Nepali Version. Int Sch Res Not 2017;2017:1–5.

20. Huang KTL, Owino C, Vreeman RC, et al. Assessment of the face validity of two pain scales in Kenya : a validation study using cognitive interviewing. BMC Palliat Care 2012;11(5):1–9.

21. Harkness J. In pursuit of quality: Issues for cross-national survey research. Int J Soc Res Methodol 1999;2(2):125–40.

22. Jowell R. How comparative is comparitive research? Am Behav Sci 1998;42(2):1998.

23. Channon AAR. Can mothers judge the size of their newborn? Assessing the determinants of a mother's perception of a baby's size at birth. J Biosoc Sci 2011;43(5):555–73.

24. Seidman DS, Slater PE, Ever-Hadani P, Gale R. Accuracy of mothers' recall of birthweight and gestational age. Br J Obstet Gynaecol 1987;94(8):731–5.

25. Hakim RB, Tielsch JM, See LC. Agreement between maternal interview- and medical record-based gestational age. Am J Epidemiol 1992;136(5):566–73.

26. Sou SC, Chen WJ, Hsieh WS, Jeng SF. Severe obstetric complications and birth characteristics in preterm or term delivery were accurately recalled by mothers. J Clin Epidemiol 2006;59(4):429–35.

27. Tomeo CA, Rich-Edwards JW, Michels KB, et al. Reproducibility and validity of maternal recall of pregnancy-related events. Epidemiology. 1999;10(6):774–7.

28. Li CY, Wei JN, Lu TH, Chuang LM, Sung FC. Mothers tended to overreport

categorical infant birth weight of their children. J Clin Epidemiol 2006;59(12):1319–25.

29. Boeke CE, Marín C, Oliveros H, Mora-Plazas M, Agudelo-Cañas S, Villamor E. Validity of maternal birthweight recall among colombian children. Matern Child Health J 2012;16(4):753–9.

30. Creswell JW. Qualitative Inquiry and Research Design. 2nd ed. Washington D.C.: Sage Publications, Inc; 2007.

4.8 Tables and Figures

Table 4.1: Length of pregnancy and birth size administered in quantitative questionnaire

Question	Answer choices
1. When your child was born, was he/she born very early, early, on time, late, or very late?	1=Very early 2=Early 3=On time 4=Late 5=Very late 9=Don’t know
Nepali: (नाम) जन्मँदा, समयभन्दा चाँडो जन्मिएको की, समयभन्दा धेरै चाँडो जन्मिएको थियो अथवा समयभन्दा ढिलो जन्मिएको की समयभन्दा धेरै ढिलो जन्मिएको थियो वा समयमै जन्मिएको थियो ? Maithili: (नाम) जनम्लैत, समयसे पहिले जनम्लै कि, समयसे बहुत पहिले जनम्लै और समय से बाद जनम्लै कि समयसे बहुत बादमे जनम्ले रहलै वा समये मे जनम्ले रहलै ?	Nepali: 1= धेरै चाँडो 2= चाँडो 3= समयमै जन्मिएको 4= ढिलो 5= धेरै ढिलो 9= थाहा नभएको
2. When your child was born, was he/she very large, larger than average, average, smaller than average, or very small?	1=Very small 2=Smaller than average 3=Average 4=Larger than average 5=Very large 9=Don’t know
Nepali: (नाम) जन्मँदा कत्रो थियो/कत्री थिइन: धेरै ठूलो, औसत भन्दा ठूलो, ठिक्कको, औसत भन्दा सानो वा धेरै सानो ? Maithili: (नाम) जनम्लै त केतना गो रहलै: बहुत बडका, साधारणसे भि बडका, ठिक ठिके, साधारण से छोट कि बहुत छोट ?	Nepali: 1= धेरै सानो 2= औसत भन्दा सानो 3= ठिक्कको 4= औसत भन्दा ठूलो 5= धेरै ठूलो 9= थाहा नभएका

Table 4.2: Characteristics of mothers interviewed in qualitative follow-up

Characteristic	n (%)
Child's age at the time of administration of the quantitative form	
<12 months	5 (42)
>=12 months	7 (58)
Child sex	
Male	7 (58)
Female	5 (42)
Place of delivery	
Home	4 (33)
Facility	8 (67)
Maternal age	
<20 yrs	4 (33)
>=20 yrs	8 (67)
Maternal Education	
No Schooling	4 (33)
Any Schooling	8 (67)
Parity	
Primiparous	6 (50)
Multiparous	6 (50)
Ethnicity	
Madhesi	12 (100)

Figure 4.1: Visual aids for birth size question

Figure 4.1: The child in photo A was 2.2kg. The child in photo B was 2.6kg. The child in photo C was 3.1kg.

Chapter 5: Validation of adjustment and imputation methods for estimation of low birthweight prevalence using a rural Nepal dataset

[Authorship to be determined]

5.1 Abstract

Background

Following recent approval by the WHO to use low birthweight (LBW) as an indicator to track health of populations, researchers and program managers now have a renewed interest on how best to produce population-based estimates of LBW at the country level. LBW estimates produced for low- and middle-income countries (LMIC) are reliant on surveys, like the Demographic and Health Survey (DHS) and Multiple Indicator Cluster Surveys (MICS). Birthweight data collected from mothers interviewed in these surveys are frequently incomplete and exhibit heavy heaping, which may result in biased estimates of LBW.

Methods

A Working Group, comprised of members from UNICEF, WHO, the London School of Hygiene and Tropical Medicine and the Johns Hopkins Bloomberg School of Public Health, have recently developed new methods to adjust estimates of LBW calculated from maternally reported information. We validated these methods using a dataset of maternal reports of birthweight against that of a "gold" standard dataset of measured

birthweights collected as part of a large community-randomized trial in rural Nepal. We also investigated patterns of missingness in the 2011 Nepal DHS dataset and systematically removed birthweights from the maternally reported dataset based on these patterns. We then imputed missing data points using variables associated with birthweight identified by the Working Group.

Findings

Current methods for adjusting LBW estimates in MICS surveys may not fully correct for underreporting of LBW. In settings where more birthweights are maternally reported rather than transcribed from a birth card, a notable degree of residual underreporting can be expected. We found the new adjustment methods developed by the Working Group resulted in more accurate LBW estimates. Applied to our birthweight dataset that exhibited heavy heaping and relied almost exclusively on maternal reports rather than birth cards, the two-component normal mixture model approach still underestimated LBW, but performed better than the existing method.

Interpretation

In a rural Nepal validation dataset with high LBW prevalence, with a large degree of heaping in primarily maternally reported birthweight, and a relatively high proportion of simulated missing birthweights, the two-component normal mixture model method generated LBW estimates more accurate than the existing method. Assessment of these methods using other validation datasets will help to understand their performance in populations with different birthweight distributions and missingness.

5.2 Introduction

Low birthweight (LBW) has long been shown to be closely associated with a greater risk of neonatal death as well as cognitive and developmental impairment and long-term health problems in adulthood.[1–3] Historically, LBW, first described in the early 1900's as birthweight <2500 grams, was used to characterize premature newborns who had an increased risk of mortality and morbidity.[3,4] Weight measurements were widely available and gestational age data was often lacking, making identification of LBW babies more feasible.[5] However, upon recognition that not all LBW babies are born early and not all babies delivered early are of LBW, the conditions, preterm and small for gestational age, have become the favored constructs in recent decades.[5] Preterm birth is defined as a birth before 37 weeks gestational age, and small for gestational age (SGA), describes newborns with birthweights below the 10th centile of a birthweight-for-gestational age reference population.[6,7] Therefore, the term LBW captures both newborns born preterm and those who are small but not necessarily preterm. While gestational age information is routinely documented in most developed countries, this remains challenging in many developing settings,[8] and as a result, LBW remains an important indicator of newborn health globally. WHO has recently approved LBW as an indicator to track health of populations, placing a renewed interest on how best to produce population-based estimates of LBW at the country level.

Nationally-representative surveys, like the Demographic and Health Surveys (DHS) and the Multiple Indicator Cluster Surveys (MICS), collect maternally reported birthweight and birth size (very large, larger than average, average, smaller than average, and very small) and generate birth indicators in low-income countries.[9] A challenge to accurately measuring LBW is that over half of newborns globally are not weighed at birth since many births still occur at home.[2,10,11] Newborns delivered in facilities may not be consistently weighed, weighed using unreliable scales, and have incomplete records.[9] In addition, mothers delivering in facilities may differ from the larger population by socioeconomic status, for example, which may bias estimates.[2,10] Mothers of newborns who were weighed at birth may be unable to produce a birthweight record or to report birthweight accurately at the time of the survey, which is sometimes administered up to five years after a birth.[9,12–15] In an analysis of DHS surveys, Blanc and Wardlaw developed a method to adjust LBW estimates to account for heaping, the tendency of numerical birthweights to be rounded to multiples of 500g in maternal reports, and missing numerical birthweights.[2,16] To address heaping at the cutoff for LBW, after removing those who weighed exactly 2500g, the percentage of newborns who weighed less than 2500g for birthweights between 2000 and 2999g was calculated in 88 DHS surveys; this averaged to 25%. Based on these results, 25% of newborns who reportedly weighed exactly 2500g were reclassified as LBW. Though some mothers may have been unable to report a numerical birthweight, they may have reported a relative birth size. To address missing birthweights, using available birth size information, the proportion of LBW newborns in each birth size category was calculated, multiplied by the overall proportion of births in each category, and summed to generate the overall number of LBW newborns. This correction is the current method used to adjust LBW estimates in MICS surveys,[16] and the adjusted estimate of LBW in Nepal in 2014 was 24.2%.[17]

A Working Group, comprised of members from UNICEF, WHO, the London School of Hygiene and Tropical Medicine and the Johns Hopkins Bloomberg School of Public Health, have recently developed new methods to adjust estimates of LBW calculated from maternally reported information. This group extended previous research describing the distribution of birthweight established across a variety of large datasets to an approach that aims to evaluate the quality of survey datasets, to identify a method using multiple imputation to handle missing birthweight, and to adjust LBW estimates accordingly. Wilcox *et al.* described birthweight as having a Gaussian distribution with a slight peak and an extended lower tail.[17] They identify two subpopulations that make up the larger distribution: a 'predominant' subpopulation with a Gaussian distribution that encompasses most birthweights and a 'residual' subpopulation made up primarily of LBW newborns.[17,18] Gage *et al.* asserted birthweight could be parametrically modeled as the sum of two Gaussian distributions and fitted both one- and two-component normal mixture models to birthweight data from different ethnic groups in New York state.[19] Charnigo *et al.* expanded on these results in his framework to propose the number of components in a normal mixture model of birthweight distribution may vary and help to identify heterogeneity in birthweight across ethnic populations.[20] Considering this prior research, the Working Group decided to include one-, two-, and three component normal mixture models in an approach to adjust LBW estimates and compare these to a crude estimate, the Blanc-Wardlaw method, and a kernel density estimation, which constructs a non-parametric curve using density estimates at each data point and a smoothing function. A first step was to assess model fit to publically available datasets of high data quality from the US (CDC National Center for Health Statistics 2015 birth data, N=3,978,497) and Mexico (Secretaria de Salud Dirreccion General de Informacion en Salud 2015 database of live birth certificates).[21,22] The Working Group established that if data quality is good, estimates resulting from each of the methods would be similar. These methods were then applied to more than 200 DHS datasets to assess data quality. The Working Group identified the two-component normal mixture model as the preferred method for generating LBW estimates.

An assessment of percent missingness in birthweight data was also completed for each DHS dataset; birthweight data were assumed to be missing at random (MAR).[23] The Working Group identified a list of 14 variables related to birthweight from a review of prior literature and in consideration of variables routinely collected in MICS and DHS surveys. A regression analysis was conducted in 88 post-2000 DHS surveys to explore the association between birthweight and the related variables; the association was averaged across the 88 DHS surveys. From these results, the Working Group reduced the list to a set of 6 variables that were overall significantly associated with birthweight: birth size (very large, larger than average, average, smaller than average, very small), sex of child (male, female), singleton/multiple births, maternal height, maternal BMI, parity (primiparous, parity 2-3, parity 4 and above). Variables were excluded from this final list either because less than a third of the 88 surveys showed a significant association to birthweight or the direction of association was inconsistent, that is variables were positively associated with birthweight in some surveys but negatively associated in others. Multiple imputation with five repetitions was performed for missing birthweights predicted by an individual's observed values in the final 6 variables in each DHS survey.

The proportion of LBW newborns was calculated using the two-component normal mixture model described above to compare estimates before and after multiple imputation. These findings and rationale for model selection will be described in detail in a separate publication.

In this paper, we validated the methods described above using a dataset of maternal reports of birthweight against that of a "gold" standard dataset of measured birthweights collected as part of a large community-randomized trial in rural Nepal. We calculated LBW estimates from maternal reports of birthweight using 6 methods: crude estimate, Blanc and Wardlaw method, non-parametric kernel density estimation, fitting a single normal curve to reported birthweight data, using a two-component normal mixture model, and fitting a three-component normal mixture model. We also investigated patterns of missingness in the 2011 Nepal DHS dataset and systematically removed birthweights from the maternally reported dataset based on these patterns. We then imputed missing data points using the 6 variables associated with birthweight identified by the Working Group. Finally, we calculated the proportion of LBW newborns using the 6 methods to compare estimates before and after imputation against that of the measured birthweight.

5.3 Methods

Study setting

The study was carried out in the rural district of Sarlahi, Nepal. This district is in terai region (plains) along the border with Bihar, India. A little more than a third of women five years and older are able to read and write.[24] Over a third of its predominantly Hindu residents are younger than 15 years of age, and about 15% of married women reported having been younger than 15 years old at the time of their first marriage.[24]

Parent trial

Conducted from November 2010 to January 2017, a randomized community-based trial examined the impact of the use of sunflower seed oil in full-body newborn massage on neonatal deaths and infections (registered at ClinicalTrials.gov (NCT01177111)). Locally-resident female project workers visited married women 15-35 years of age at home every 5 weeks to identify new pregnancies; pregnancies among women outside this age range were identified informally. All pregnant women agreeing to participate in the parent trial were asked to report the number of their previous live births, and the median of three measures of the women's height and weight were recorded. Enrolled women were followed through delivery; study staff visited as soon as possible after delivery and through the first month (days 1, 3, 7, 10, 14, 21, and 28). At the first visit, workers recorded date/time of delivery, circumstances of labor and delivery, health status of mother and newborn, child's sex, whether it was a single or multiple birth, and the median of three measures of the baby's weight using a digital scale precise to 10g (Tanita BD-585). The date, time of birth and weight of the newborn were also provided to the mother/caretaker on a small 10 x 8 centimeter card. Subsequent visits focused on maternal report and directly observed aspects of newborn health.

Substudy

We selected mother/child pairs from the parent trial for one additional follow-up visit to ask mothers to report on labor and delivery, immediate newborn care, postnatal care, and neonatal morbidity and care seeking in the first 7 days of life. We aimed to interview roughly the same numbers of mothers at each follow-up time at 1, 3, 6, 9, 12, 18, or 24 months after birth. Study staff requested participation in the homes of selected mothers, administered oral consent in Nepali or Maithili (a local language) and obtained a signature or thumbprint for those who agreed to participate. Specific to this analysis, mothers were asked about birthweight, birth size, and whether they had a written record of their child's birthweight.

Ethical approval

The parent trial and the substudy both received ethical approval from the Johns Hopkins Bloomberg School of Public Health Institutional Review Board, Baltimore, MD, USA. Local approval was received from the Tribhuvan University Institute of Medicine, Kathmandu, Nepal for the parent trial and from the Nepal Health Research Council, Kathmandu, Nepal for the substudy.

Methods for estimating LBW using maternally reported birthweight

Stata version 14.0 (StataCorp, College Station, TX, USA) was used for this analysis. We compared the proportion of LBW, preterm birth and SGA newborns in the substudy and the parent trial. Using the eligibility criteria of the substudy, those in the parent trial were restricted to live singleton births who had a first household visit completed within 72 hours after birth. SGA was calculated using the INTERGROWTH-21 standard growth curve.[25,26] Six methods were applied to maternally reported birthweights to estimate LBW. The first generated a crude unadjusted estimate of the proportion of birthweights less than 2500g. The second used the Blanc and Wardlaw two-part adjustment procedure in which i) 25% of births reported as exactly 2500g were reclassified as LBW and ii) the proportion of LBW within each birth size category based on newborns who were weighed was multiplied by the total number of births to estimate the number of LBW newborns among all births.[16] Third, a kernel density estimation method was used to construct a non-parametric curve using density estimates at each data point and a smoothing function across our dataset of maternally reported birthweights.[27] Setting a threshold at 2500g, the proportion of LBW newborns was calculated as the area under the kernel density plot. A fourth method fit a normal curve to the data with a mean and standard deviation, and calculated the LBW proportion as the area under the curve below 2500g. A fifth method utilized a two-component normal mixture model in which the distribution of birthweights was assumed to be composed of two subpopulations: i) a primary normal distribution that accounts for most birthweights, and ii) a secondary normal distribution that captures the smallest newborns in the left tail of the distribution.[19] Combining these two curves, the area under the overall function with a cut point at 2500g generated the proportion of LBW newborns. Finally, a three-component normal mixture model generated normal curves fitted to most birthweights around the mean, the smallest newborns, and an additional curve for the largest newborns in the right tail of the distribution.[20] The area under the overall mixture curve below 2500g produced the proportion of LBW newborns.

Patterns in missingness

We aimed to produce datasets with missing birthweights based on patterns of missingness observed in the 2011 Nepal DHS dataset. We hypothesized *a priori* that missing birthweight may be associated with whether the mother had at least four antenatal care (ANC) visits, birth order of the child, birth size, child sex, singleton/multiple, maternal height, maternal BMI, maternal smoking status, birth interval, maternal education, maternal age, and wealth quintile. We first restricted the 2011 Nepal DHS dataset to only households in rural areas to ensure comparability to our study. The wealth quintile variable was reconstructed to reflect the distribution of wealth among rural households. We conducted a logistic regression using individual sampling weights adjusting for differential probability in selection to investigate the association of missing birthweight and the above variables.[28,29]

Multiple Imputation

We first removed birthweights from our dataset of maternal reports at random. Based on the results of the patterns of missingness among rural households from the 2011 Nepal

DHS dataset, we also removed birthweights in this dataset for each variable that was significantly associated with missing birthweight. We noted the percent missing birthweights within each category of these variables, and removed birthweights at random within each category in our dataset. For each of these datasets with artificially missing birthweights, we conducted multiple imputation with five repetitions using the variables associated with birthweight identified by the Working Group. These variables included: birth size, sex of child, maternal height, maternal BMI, parity; singleton/multiple birth was not used since multiple births was an exclusion criteria for our substudy. Finally, we compare these estimates to those generated without imputation, those from our complete dataset of maternally reported birthweight, and those from our dataset of measured birthweight.

5.4 Results

1528 mothers consented and were interviewed (**Figure 5.1**). After excluding 15 participants (birth assessment >72 hours after birth [n=3], twin delivery [n=1], repeat participation [n=11]), a total of 1513 mother/child pairs were included. Of these, 16 (1.1%) children were missing a digital weight measurement for reasons that included deaths prior to measurement (n=14), parental refusal of weight measurement (n=1), and missing weight measurement (n=1). 74 (4.9%) children were missing maternally reported birthweights, where mothers reported the child was not weighed (n=21), was uncertain if child was weighed (n=6), or was weighed but could not provide a numerical weight (n=47). Of the 1486 mothers who were asked if they had a card with a birthweight record, only 22 (1.5%) presented cards provided by a facility. To mimic circumstances of a DHS or MICS survey, we used birthweights recorded on these facility cards in our dataset of maternally reported birthweights. **Table 5.1** presents the proportion of LBW, preterm birth, and SGA newborns in the substudy compared to the parent trial.

Figure 5.2 displays the kernel density and one-component normal curves fitted to the measured and reported birthweight datasets. The measured birthweights appear to be generally normally distributed with a slightly higher peak and a left tail that diverges somewhat from the normal curve. Heavy heaping at 2000, 2500, 3000, and 3500g is evident in the kernel density curve fitted to the reported birthweight dataset. **Table 5.2** presents the means and standard deviations (SD) of normal curves fitted to both measured and reported birthweights. The mean and SD of the reported birthweight were both higher than that of the measured birthweight. **Figure 5.3** presents the two- and three-component normal curves overlaying the kernel density and one-component normal curves for comparison for reported birthweights. Means and SDs are also presented in **Table 5.2** for the two- and three-component normal curves. The methods appear to better fit normal curves to the measured birthweights than the reported birthweights. In both the two- and three-component normal mixture models for the reported birthweight dataset, the second and third components, respectively, appear flat and the proportions of LBW under those curves are 0 but they still contribute to the shape of the mixture curves.

Figure 5.4 presents the LBW point estimates from both the measured and reported birthweight datasets using 6 methods: crude estimate, Blanc and Wardlaw method, non-parametric kernel density estimation, fitting a single normal curve to reported birthweight data, using a two-component normal mixture model, and fitting a three-component normal mixture model. LBW estimates calculated from measured birthweights are relatively similar, ranging from a crude estimate of 27.7% to 30.3% when fitting a normal curve. Estimates using reported birthweights showed much more variation from a crude estimate of 17.1% to 26.7% calculated from a kernel density estimation. Crude estimates show a 10% absolute difference in LBW comparing measured and reported datasets. Assuming the measured birthweight is of good data quality ("gold" standard) as evidenced from its fairly normal distribution, adjusting the reported birthweight data using the kernel density (26.7%) and two-component normal mixture model methods (26.4%) appear to generate LBW point estimates closest to our "true" measured data.

From our analysis of the 2011 Nepal DHS rural household dataset, 2866 (68%) had missing birthweights. The odds of missing birthweight was significantly associated with birth size, single versus multiple births, parity, having at least four ANC visits, rural wealth quintile, and maternal education (**Table 5.3**). **Table 5.4** presents the percent missing birthweight within each of the categories of variables significantly associated with missing birthweight using Nepal DHS 2011. Single versus multiple birth was excluded since this was an exclusion criteria in our substudy. A total of 6 datasets with artificially missing birthweights were created: one with 68% of birthweights removed at random and five more with birthweights systematically removed following patterns of missingness observed in the Nepal DHS 2011 by birth size, parity, ANC status, rural wealth quintile, and maternal education.

After multiply imputing missing birthweights, we a fit kernel density and one-, two-, and three-component curves. **Figure 5.5** presents an example of these curves fitted to a first imputation of missing birthweights following removal of 68% of birthweights by wealth quintile. Comparing these kernel density curves to those of the reported birthweights prior to removal and multiply imputing birthweights, the kernel density curves are smoother; however, heaping is still evident in most, especially at 2500 and 3000g. **Table 5.5** summarizes the LBW point estimates calculated using the 6 methods after multiple imputation in each of the datasets with artificially missing birthweights. Across all datasets, after multiple imputation, all crude estimates are higher than those calculated from the reported birthweight dataset (17.1%), ranging from 20.8% in the dataset that removed birthweight by patterns of missingness in the ANC status variable to 24.8% in those created by missing patterns in parity and rural wealth quintile. We observed a similar pattern in estimates calculated using the Blanc-Wardlaw method post-multiple imputation; however, all LBW estimates remained lower than that of our "gold" standard (27.7%). Using kernel density estimation, estimates are close to our "gold" standard, ranging from 25.7% in the dataset where birthweight had been removed by patterns of missingness in the maternal education variable to 27.9% in that of the rural wealth quintile variable. LBW estimates calculated from fitting a normal curve to distributions of multiply imputed birthweights were higher than that of our "gold" standard in datasets, where birthweights had been removed by birth size (28.6%) and by rural wealth quintile (28.2%). From the two-component normal mixture model, estimates were lower than that of the single normal curve, ranging from 25.2% in the dataset, where birthweights had been removed by maternal education, to 27.5%, where birthweights had been removed by rural wealth quintile. Finally, comparable or slightly higher LBW estimates were generated from the three-component normal mixture model compared to those of the two-component normal mixture model. The two highest estimates (27.7% and 27.5%) were calculated from the datasets in which birthweights had been removed by birth size and rural wealth quintile, respectively.

5.5 Discussion

The current methods for adjusting LBW estimates in MICS surveys, as developed by Blanc and Wardlaw,[2,16] may not fully correct for biased reporting of LBW in surveys. Developed from an analysis of 88 DHS surveys, these methods apply an adjustment for heaped birthweights developed from an averaged pattern across surveys. However, heaping can be highly variable and the resulting LBW estimate is particularly sensitive to this variation.[30] Channon *et al.* reported greater heaping in maternally recalled birthweights compared to those from a birth card.[30] In countries where the majority of birthweights were recalled from memory versus recorded from a birth card, the study also found a significantly higher mean birthweight in the former compared to the latter.[30] Therefore, in settings where more birthweights are maternally reported rather than transcribed from a birth card, a notable degree of residual underreporting after applying the Blanc-Wardlaw adjustment can be expected, as observed in our study. Additionally, the method relies on consistent perception of birth size. Studies have assessed the relationship of birthweight and birth size within DHS datasets and found that mean birthweight generally decreased with decreasing birth size, consistent with our findings (results not shown here; reference publication).[12,13,15,31] However, maternal perception of birth size may vary across different populations. As Channon describes, mothers' perception of birth size may be affected by various neighborhood and regional factors specific to a setting that shape a reference for how mothers assess their child's size.[32] As we have shown previously, this phenomenon applies in this rural Nepal population, where mothers perceive smaller children as being of average size (**Chapter 4**).

Based on previous efforts by the Working Group to develop new approaches for adjustments, we found that these may result in more accurate LBW estimates. These methods make use of prior studies that have explored the distribution of birthweight and its association with the distribution of neonatal mortality in large, complete datasets from a variety of subpopulations in developed countries.[17–20,33] After examining the performance of these methods across more than 200 DHS datasets, the Working Group identified the two-component normal mixture model as the preferred method, considering accuracy and parsimony. Our birthweight dataset exhibited heavy heaping and relied almost exclusively on maternal reports rather than birth cards. This approach still underestimated LBW, but performed better than the Blanc-Wardlaw method.

We identified factors associated with missing birthweight in the 2011 Nepal DHS dataset that were similar to findings in prior studies, including lower parental education, lower socioeconomic status, and higher parity.[12,16,30,34–36] We were unable to look at missing birthweight by urban versus rural residence since our study site included only rural households. Mothers who delivered in a facility are also more likely to have had at least four ANC visits and less likely to have a missing birthweight.[35] Similar characteristics are also associated with a higher birthweight.[37] From our exercise in creating datasets that attempted to mimic selection bias in birthweight missingness patterns in DHS and MICS surveys followed by multiple imputation, the two-component normal mixture model calculated a LBW estimate closest to our "gold" standard in the dataset with birthweights that had been removed by wealth quintile. This may mean that this method provides more accurate adjustments when birthweight missingness is closely related to lower household wealth status. In reality, however, selection bias in birthweight missingness is likely related to a combination of different factors acting simultaneously, limiting the applicability of our exercise. In addition, our exercise used only variables available in both the DHS and our study; we were unable to include variables like single versus multiple births and other possible factors, such as rural versus urban environments, that may result in selection bias.

A strength of our study is the use of accurate and calibrated scales of research quality to minimize bias in measurement that likely operated at least as well as those used in the large datasets in prior studies used to develop normal mixture methods describing the birthweight distribution.[17,19,20] The high quality of our measurement data is evidenced by the similar distribution pattern of birthweights noted in these papers.[17,19,20] A limitation of using these measurements as our "gold" standard is that newborns were weighed up to 72 hours after birth. In these first hours of life, newborns generally lose weight before growth and weight gain are observed. Therefore, our measurements likely overestimate

LBW; however, our intention was to validate these methods rather than provide an estimate of prevalence. Additionally, in the case of home births, mothers were likely recalling the birthweight measured during the parent trial since this would have been the only birthweight provided to them. However, for those who delivered in a facility, their children may have been weighed both at the facility and during participation in the parent trial. In the latter case, we assumed the mother was reporting the weight measurement provided to them during the parent trial. Finally, our validation of these methods may have limited generalizability as our study population had a relatively high proportion of LBW newborns, and our normal mixture curves are shifted down towards lower birthweights. We also removed a relatively high percentage of birthweights to mimic the patterns among rural households in the 2011 Nepal DHS dataset. Additionally, considering that the majority of birthweights in our dataset were recalled by mothers rather than transcribed from birth cards, which resulted in heavy heaping, our dataset might represent a fairly extreme case of birthweights that would require adjustment. Therefore, it may be that if these methods performed adequately in this dataset, they may be sufficient in datasets with lower frequencies of missing birthweight and a lower degree of heaping resulting from less reporting bias.

5.6 Conclusion

Existing methods to adjust LBW estimates to address heaping, misclassification, and missing birthweights in MICS surveys may be insufficient and result in residual underreporting of LBW. New methods developed by a Working Group show promise in producing more accurate LBW estimates. Applied to a rural Nepal validation dataset with high LBW prevalence, a large degree of heaping with birthweights primarily from maternal recall, and in simulations, a relatively high proportion of missing birthweights, the two-component normal mixture model method generated LBW estimates more accurate than the Blanc-Wardlaw method. Assessment of these methods using other validation datasets will help to understand their performance in populations with different birthweight distributions. In future validation studies, a tailored investigation of patterns in missing birthweight is required for each DHS or MICS dataset to be applied to potential validation datasets. In terms of program application, these methods are more complex than existing methods, and program managers conducting DHS and MICS surveys may need to weigh the benefits of improved accuracy against increased complexity.

5.7 References

1. Unicef. Committing to Child Survival : A Promise Renewed. Unicef Progress Report 2014. 2014.

2. United Nations Children's Fund and World Health Organization. Low Birthweight: Country, regional and global estimates. 2004.

3. Wilcox AJ, Russell IT. Birthweight and Perinatal Mortality: II. On Weight-Specific Mortality. Int J Epidemiol 1983;12(3):319–25.

4. Salihu HM. The historic origin of Prematurity/Low birth weight: Lessons from the past. Available from: Am Acad of Pediatrics n.d.

5. Wilcox a J. On the importance--and the unimportance--of birthweight. Int J Epidemiol 2001;30(6):1233–41.

6. The World Health Organization (WHO). Physical Status: the Use and Interpretation of Anthropometry. Report of a WHO Expert Committee. Technical Report Series No. 854. 1995.

7. Lee AC, Katz J, Blencowe H, et al. National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. Lancet Global Heal 2013;1(1):26-36.

8. March of Dimes, PMNCHS, Save the Children, WHO. Born Too Soon: The Global Action Report on Preterm Birth. Eds CP Howson, MV Kinney, JE Lawn. Geneva: 2012.

9. Hancioglu A, Arnold F. Measuring Coverage in MNCH: Tracking Progress in Health for Women and Children Using DHS and MICS Household Surveys. PLoS Med 2013;10(5).

10. United Nations Children's Fund (UNICEF). Improving child nutrition. The achievable imperative for global progress. New York, NY: 2013.

11. WHO. Global Nutrition Targets 2025: Low Birth Weight Policy Brief. Geneva, Switzerland: 2014.

12. Boerma JT, Weinstein KI, Rutstein SO, Sommerfelt AE. Data on birth weight in developing countries: Can surveys help? Bull World Health Organ 1996;74(2):209–16.

13. Eggleston E, Tsui AO, Fortney J. Assessing survey measures of infant birth weight and birth size in Ecuador. J Biosoc Sci 2000;32(3):373–82.

14. Lule S a., Webb EL, Ndibazza J, et al. Maternal recall of birthweight and birth size in Entebbe, Uganda. Trop Med Int Heal 2012;17(12):1465–9.

15. Mbuagbaw L, Gofin R. Can recall of birth size be used as a measure of birthweight

in Cameroon? Paediatr Perinat Epidemiol 2010;24(4):383–9.

16. Blanc AK, Wardlaw T. Monitoring low birth weight: An evaluation of international estimates and an updated estimation procedure. Bull World Health Organ 2005;83(3):178–85.

17. Central Bureau of Statistics. Nepal Multiple Indicator Cluster Survey. Kathmandu, Nepal: 2015.

18. Wilcox AJ, Russellf IT. Birthweight and Perinatal Mortality: I. On the Frequency Distribution of Birthweight. Int J Epidemiol 1983;12(3):314–8.

19. Umbach DM, Wilcox AJ. A technique for measuring epidemiologically useful features of birthweight distributions. Stat Med 1996;15:1333–48.

20. Gage TB, Therriault G. Variability of Birth-Weight Distributions by Sex and Ethnicity: Analysis Using Mixture Models. Hum Biol 1998;70(3):517–34.

21. Charnigo R, Chesnut LW, Lobianco T, Kirby RS. Thinking outside the curve, part I: modeling birthweight distribution. BMC Pregnancy Childbirth 2010;10:37.

22. CDC National Center for Health Statistics. Birthweight and Gestation. 2017 [cited 2017 Apr 11];Available from: https://www.cdc.gov/nchs/fastats/birthweight.htm

23. Secretaria de Salud Dirreccion General de Informacion en Salud. Databases of Live Birth Certificates. 2016 [cited 2017 Apr 11]; Available from: http://www.dgis.salud.gob.mx/contenidos/basesdedatos/std_nacimientos_gobmx.html

24. Schafer JL, Graham JW. Missing data: Our view of the state of the art. Psychol Methods 2002;7(2):147–77.

25. Central Bureau Of Statistics/Government of Nepal. National Population and Housing Census 2011. Kathmandu, Nepal: 2012. Available from: http://cbs.gov.np/?p=2017

26. Villar J, Ismail LC, Victora CG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. Lancet 2014;384(9946):857–68.

27. INTERGROWTH-21st. The International Fetal and Newborn Growth Consortium for the 21st Century. 2017 [cited 2017 May 15];Available from: https://intergrowth21.tghn.org/

28. Kerm P Van. Adaptive kernel density estimation. In: 9th UK Stata Users meeting, Royal Statistical Society. London, UK: 2003.

29. Rutstein SO, Rojas G. Guide to DHS statistics. Demographic and health surveys methodology. Calverton, Maryland: 2006.

30. Ministry of Health and Population (MOHP) [Nepal], New ERA, ICF International Inc. Nepal Demographic and Health Survey 2011. Kathmandu, Nepal: 2012.

31. Channon AAR, Padmadas SS, McDonald JW. Measuring birth weight in developing countries: Does the method of reporting in retrospective surveys matter? Matern Child Health J 2011;15(1):12–8.

32. Islam MM. Can maternal recalled birth size be used as a proxy measure of birth weight? An evaluation based on a population health survey in Oman. Matern Child Health J 2014;18(6):1462–70.

33. Channon AAR. Can mothers judge the size of their newborn? Assessing the determinants of a mother's perception of a baby's size at birth. J Biosoc Sci 2011;43(5):555–73.

34. Gage TB. Birth-Weight-Specific Infant and Neonatal Mortality: Effects of Heterogeneity in the Birth Cohort. Hum Biol 2002;74(2):165–84.

35. Moreno L, Goldman N. An assessment of survey data on birthweight. Soc Sci Med 1990;31(4):491–500.

36. Subramanyam MA, Ackerson LK, Subramanian S V. Patterning in birthweight in India: Analysis of maternal recall and health card data. PLoS One 2010;5(7):1–9.

37. Robles A, Goldman N. Can accurate data on birthweight be obtained from health interview surveys? Int J Epidemiol 1999;28(5):925–31.

38. Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. Bull World Health Organ 1987;65(5):663–737.

5.8 Tables and Figures

Figure 5.1: Flowchart for participant selection

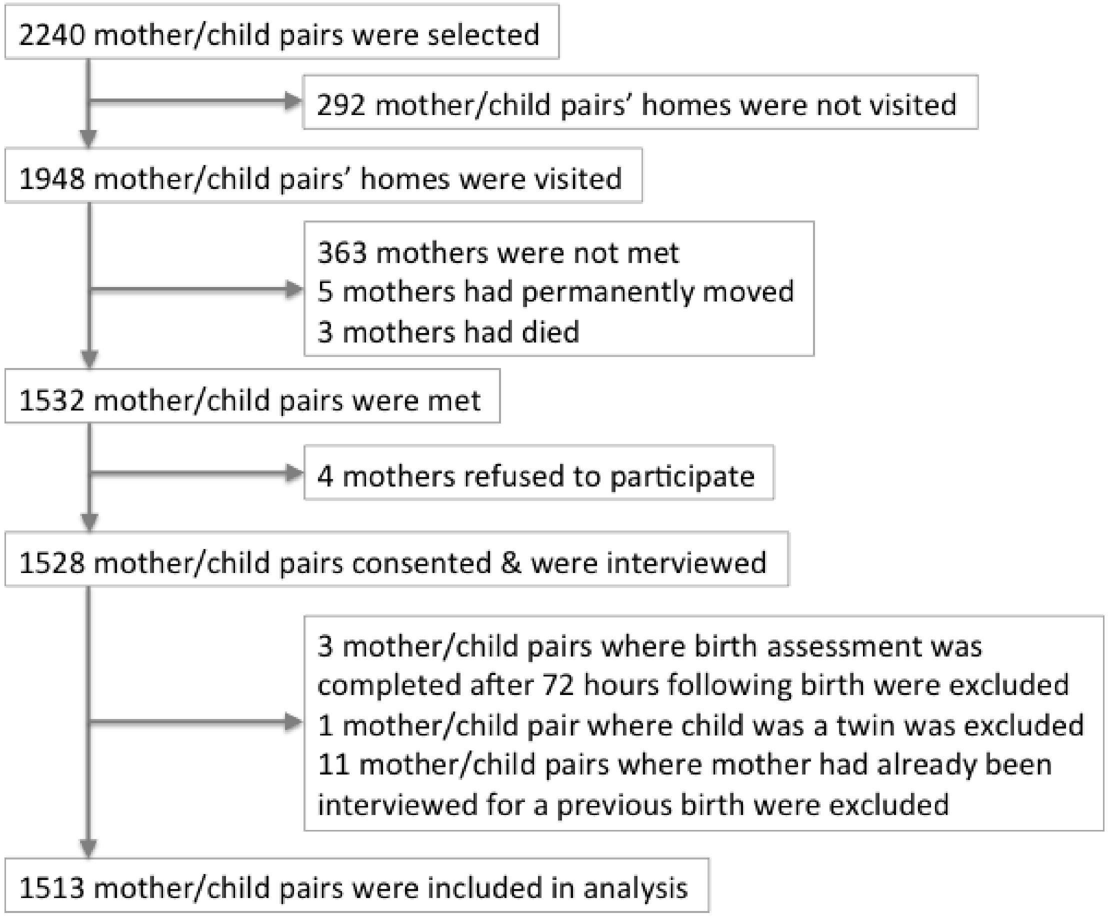

Table 5.1: Proportion of low birthweight, preterm birth and small-for-gestational-age in substudy and parent trial

	Substudy		Parent Trial*	
	N	n (%)	N	n (%)
LBW	1497	414 (27.7)	21842	6534 (29.9)
Preterm birth	1507	243 (16.1)	21946	3368 (15.4)
SGA#	1351	594 (44.0)	21317	9862 (46.3)

Table 5.1: Proportion of low birthweight, preterm birth and small-for-gestational-age

*Newborns in the parent trial were restricted to live births, singletons, and those who had a first household visit completed within 72 hours after birth, comparable to the eligibility criteria of the substudy.

#Small-for-gestational-age (SGA) defined as weight below the 10th percentile for the gestational age and sex using the INTERGROWTH-21 standard growth curve

Figure 5.2: Kernel density curve and one-component normal curves fitted to the measured and reported birthweight datasets

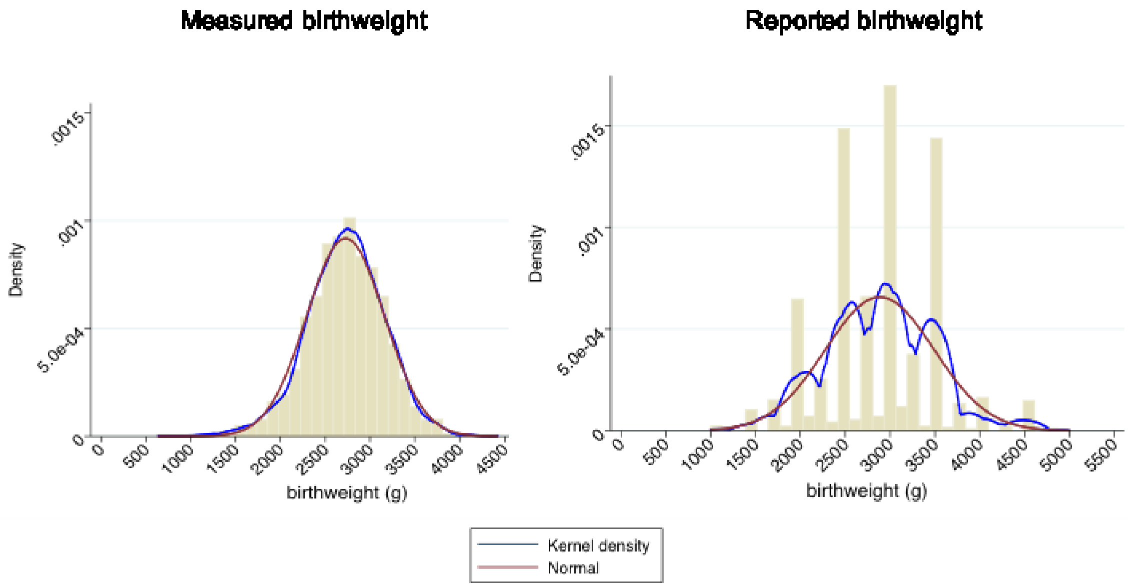

Table 5.2: Normal mixture curves fitted to measured and reported birthweights

	Mean (SD) (g)
Measured	
Single normal curve	2726 (434)
Two-component normal mixture	
Normal component 1	2531 (643)
Normal component 2	2758 (381)
Three-component normal mixture	
Normal component 1	2567 (601)
Normal component 2	2581 (293)
Normal component 3	2985 (308)
Reported	
Single normal curve	2885 (607)
Two-component normal mixture	
Normal component 1	2854 (573)
Normal component 2	4500 (1.1E-07)
Three-component normal mixture	
Normal component 1	2854 (569)
Normal component 2	4500 (2.8E-07)
Normal component 3	2997 (446)

Figure 5.3: Kernel density curve, two-, and three-component normal curves fitted to the reported birthweight dataset

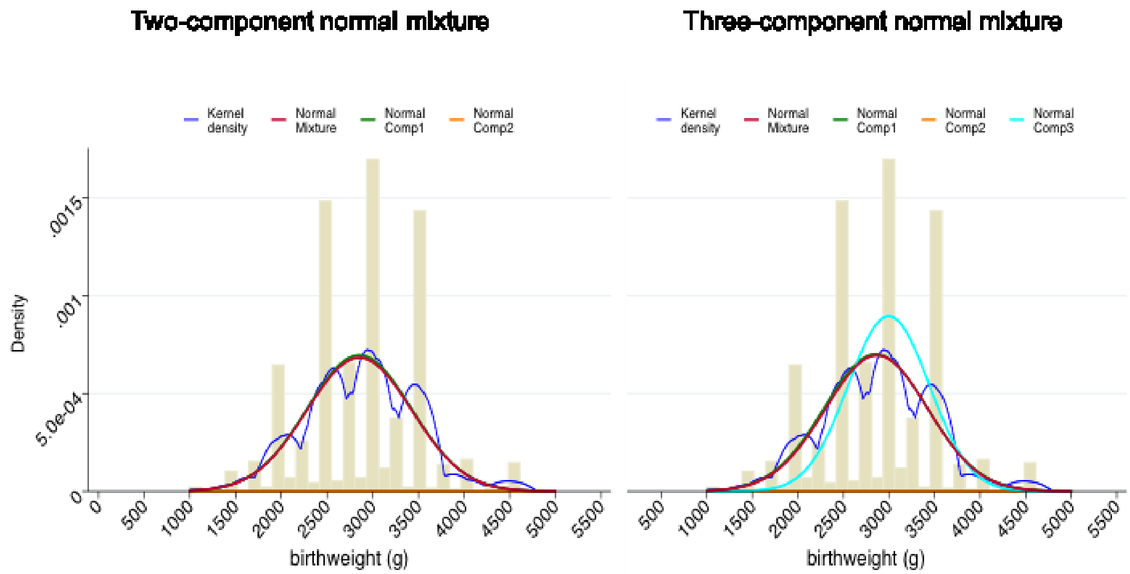

Figure 5.4: LBW estimates generated from six methods using measured birthweight dataset

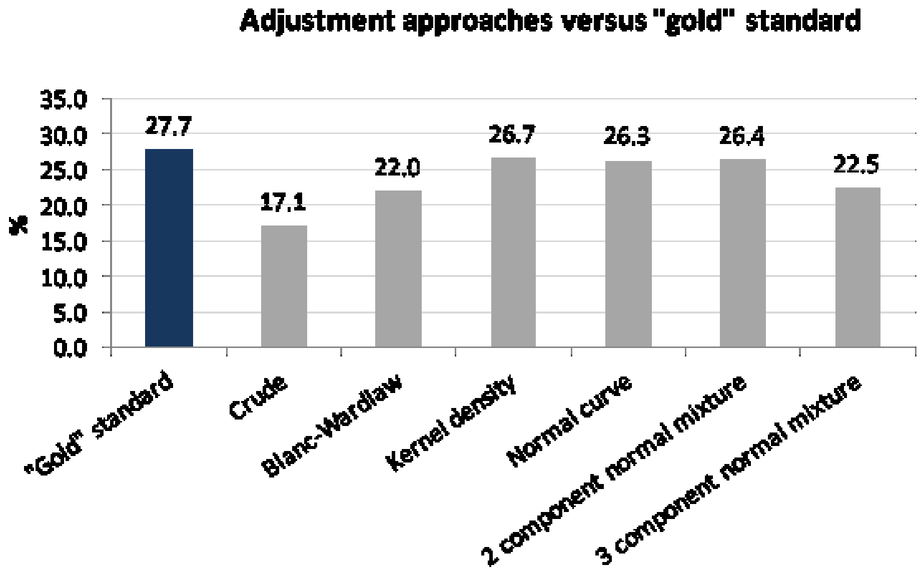

Table 5.3: Unadjusted and adjusted odds ratios (OR) for missing birthweight by socioeconomic and health risk factors among rural households in the 2011 Nepal DHS survey

	n	Unadjusted OR (95% CI)	p-value	n	Adjusted OR (95% CI)	p-value
Birth size						
Very large (ref)	80			32		
Larger than average	686	0.93 (0.51, 1.71)	0.82	257	1.30 (0.50, 3.37)	0.58
Average	2666	1.03 (0.57, 1.86)	0.91	949	1.80 (0.72, 4.48)	0.21
Smaller than average	601	1.56 (0.83, 2.91)	0.17	236	2.93 (1.10, 7.79)	0.03
Very small	178	1.35 (0.67, 2.73)	0.40	63	2.06 (0.66, 6.40)	0.21
Child sex						
Male (ref)	2191			811		
Female	2024	1.03 (0.88, 1.21)	0.68	726	0.93 (0.70, 1.22)	0.59
Single vs Multiple						
Singleton (ref)	4159			1528		
Multiple births	56	0.37 (0.20, 0.71)	0.003	9	0.21 (0.04, 1.08)	0.06
Maternal height (cm)	2034	0.97 (0.95, 0.99)	0.007	1537	0.99 (0.97, 1.02)	0.69
Maternal BMI (kg/m2)	2033	0.96 (0.92, 1.00)	0.08	1537	1.06 (1.01, 1.12)	0.02
Maternal smoking status						
Non smoker (ref)	3483			1289		
Smoker	732	3.31 (2.49, 4.41)	<0.001	248	1.02 (0.62, 1.68)	0.94
Parity						
Primiparous (ref)	1314			472		
Second or third child	1842	2.38 (2.00, 2.84)	<0.001	676	2.02 (1.29, 3.16)	0.002
Fourth child or greater	1059	5.11 (3.99, 6.57)	<0.001	389	2.86 (1.60, 5.09)	<0.001
ANC status						
Less than 4 visits (ref)	1646			811		
4 visits or more	1536	0.17 (0.14, 0.21)	<0.001	726	0.29 (0.22, 0.40)	<0.001

	n	Unadjusted OR (95% CI)	p-value	n	Adjusted OR (95% CI)	p-value
Rural wealth quintile						
WQ 1 (ref)	1436			465		
WQ 2	976	0.47 (0.36, 0.61)	<0.001	358	0.51 (0.32, 0.79)	0.003
WQ 3	752	0.28 (0.21, 0.37)	<0.001	298	0.40 (0.25, 0.64)	<0.001
WQ 4	579	0.16 (0.12, 0.22)	<0.001	232	0.24 (0.15, 0.39)	<0.001
WQ 5	472	0.06 (0.04, 0.08)	<0.001	184	0.08 (0.04, 0.14)	<0.001
Birth interval (mos)	4212	1.01 (1.00, 1.01)	<0.001	1537	1.00 (0.99, 1.01)	0.55
Maternal education						
none (ref)	2157			742		
any primary	864	0.47 (0.38, 0.59)	<0.001	316	0.92 (0.62, 1.36)	0.68
any secondary or higher	1194	0.16 (0.13, 0.20)	<0.001	479	0.53 (0.36, 0.78)	0.001
Maternal age (yrs)						
15-19	850			282		
20-34	3032	1.56 (1.29, 1.89)	<0.001	1120	1.34 (0.92, 1.94)	0.12
35 and older	329	3.00 (2.02, 4.45)	<0.001	135	0.97 (0.44, 2.14)	0.95

Table 5.4: Missing birthweight for each significantly associated variable among rural households in the 2011 Nepal DHS dataset

	% missing
Birth size	
Very large	63.7
Larger than average	60.9
Average	67.5
Smaller than average	76.1
Very small	65.1
Parity	
Primiparous	49.6
Second or third child	70.4
Fourth child or greater	87.0
ANC status	
Less than 4 visits	82.6
4 visits or more	43.4
Rural wealth quintile	
WQ 1	88.6
WQ 2	78.5
WQ 3	68.3
WQ 4	55.1
WQ 5	30.4
Maternal education	
none	82.2
any primary	70.8
any secondary or higher	40.9

Figure 5.5: Kernel density curve, one-, two-, and three-component normal curves fitted to a first imputation of missing birthweights after removing birthweights by rural wealth quintile

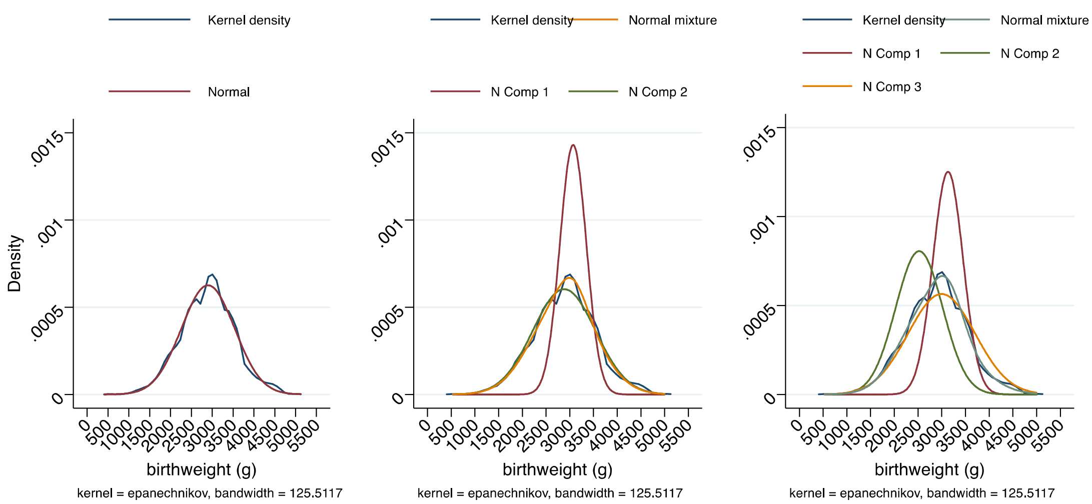

Table 5.5: LBW point estimates (%) calculated using 6 methods in multiply imputed datasets (r=5) after removing birthweights at random and systematically

Pattern of birthweight removal	Crude	Blanc-Wardlaw	Kernel density	Normal curve	2 component normal mixture	3 component normal mixture
At random	23.2	24.7	26.3	26.4	25.8	25.9
By birth size	24.3	25.9	27.6	28.6	27.4	27.7
By parity	24.8	26.0	26.9	26.9	26.5	26.9
By ANC status	20.8	23.3	25.8	25.7	25.4	26.3
By rural wealth quintile	24.8	26.3	27.9	28.2	27.5	27.5
By maternal education	23.2	24.4	25.7	25.5	25.2	25.3

Chapter 6: Discussion

6.1 Summary of results

Nepal has seen dramatic reductions in child and neonatal mortality in recent years; however, more still needs to be done in order to reach the SDGs goals by 2030. To track global progress towards meeting these targets, methods used in household surveys need to be validated in a variety of settings. Accurate LBW and preterm birth estimates can help us to identify newborns most in need of interventions and serve as denominators of measures of coverage for interventions aimed at further reducing mortality rates. Findings presented in this thesis described the validity of these indicators in a rural Nepali setting and possible strategies to improve estimates in data collection and in analysis.

Objective 1: Validity of LBW and preterm birth indicators in rural Nepal

The purpose of this objective was to validate the LBW and preterm birth indicators calculated from maternally reported birthweight, birth size and length of pregnancy, and to assess whether the accuracy of maternal reports diminished with increasing time since birth. We observed low individual-level accuracy (AUC=0.69) and high population-level bias (IF=0.62) for the LBW indicator using maternally reported birthweight in grams. LBW using maternally reported birth size and the preterm birth indicator had lower individual-level accuracy (AUC=0.58, 0.56, respectively) and greater degree of population-level bias (IF=0.28, 0.35, respectively) up to 24 months following birth. Length of recall time did not affect accuracy of LBW indicators when controlling for other possible confounders. For the preterm birth indicator, length of recall time was not significantly associated with accuracy up to 20 months after birth; however, after 20 months, accuracy statistically significantly improved although the association was only slightly great than null, adjusting for other confounders.

Objective 2: Maternal understanding of LBW and preterm birth questions in rural Nepal

The aim of this objective was to investigate maternal comprehension of administered questions used to ascertain relative birth size and length of gestation, and to identify potential visual aids that may assist in scaling these questions within a cultural frame of reference to inform interpretation of data collected from maternal reports. FGDs with study staff identified challenges in the translation of the birth size and length of pregnancy questions. In IDIs with mothers, we described the potential utility of using photographs of children of different sizes to aid our understanding of the local perception of birth size. From both study staff and mothers, a longer period of time since the birth of a child was thought to be associated with diminished accuracy of maternal reports, a perception consistent with our initial hypothesis but not supported by our quantitative findings. We also described the relationship dynamics between study staff and mothers who participated and how this may have affected mothers' responses in our quantitative study.

Objective 3: Validating methods to multiply impute missing birthweights and adjust the LBW indicator using a rural Nepal dataset

The goal of this objective was to validate methods developed by a Working Group, comprised of members from UNICEF, WHO, the London School of Hygiene and Tropical Medicine and the Johns Hopkins Bloomberg School of Public Health to adjust estimates of LBW calculated from maternally reported information, which is frequently incomplete and exhibits heavy heaping. We also investigated patterns of missingness in the 2011 Nepal DHS dataset and systematically removed birthweights from the maternally reported dataset based on these patterns. We then imputed missing data points using variables associated with birthweight identified by the Working Group. Current methods for adjusting LBW estimates in MICS surveys may not fully correct for underreporting of LBW. In settings where more birthweights are maternally reported rather than transcribed from a birth card, we expected a notable degree of residual underreporting of LBW. We found the new adjustment methods developed by the Working Group may result in more accurate LBW estimates. Applied to our birthweight dataset that exhibited heavy heaping and relied almost exclusively on maternal reports rather than birth cards, a two-component normal mixture model approach to adjust this reporting error still underestimated LBW, but performed better than the existing method.

6.2 Strengths and limitations

Considering much of the prior validation literature included only facility births, one strength of our study is the inclusion of mothers who delivered in the home. In the South Asian region, where around 69% of newborns are not weighed at birth, perhaps mothers place less importance on remembering and documenting birthweight, as evidenced by very few mothers who were able to present a birthweight card. Our population was characterized by a relatively high prevalence of LBW and preterm birth, allowing us to demonstrate that these indicators may be increasingly vulnerable to being underestimated in populations with higher prevalences. Another strength was the use of accurate and calibrated infant weighing scales of research quality and trained and supervised data collectors, in contrast to many delivery facilities, for the "gold" standard. The high quality of these measurements is evidenced by its relatively Gaussian distribution when overlaying a normal curve onto the kernel density estimate.

A limitation of our birthweight measurements used as the gold standard is that newborns were weighed up to 72 hours after birth. In this time period, newborns normally lose weight before patterns of growth and weight gain are observed. Therefore, our measurements were likely taken at a nadir and overestimate the prevalence of LBW; however, our intention was to validate maternal report rather than provide an estimate of prevalence. In addition, for home births, we are fairly confident that mothers were reporting the birthweight measured during the parent trial since this would have been the only birthweight provided to them. However, mothers who delivered in a facility may have had their child weighed both at the facility and during participation in the parent trial. For facility births, we assumed the mother was reporting the weight measurement during the parent trial. We did not ask mothers to report birthweight immediately after the measurement was taken, which would have provided more information about whether mothers could retain birthweight information if the event occurred just prior to our interview.

This thesis research only asked mothers to report a relative length of pregnancy and did not ask them to report a numerical gestational age. This limited our conclusions regarding the accuracy of maternal reports for classifying preterm births. However, this was an attempt (although unsuccessful) to estimate preterm prevalence in low income country settings by asking questions about length of pregnancy where ultrasound is not available for gestational age dating and recall of dates of last menstrual period are poor. In addition, we adapted the phrasing of the relative length of pregnancy question from the relative birth size question in the DHS and MICS surveys; however, the length of pregnancy question had not been used elsewhere. The question had also been translated from English to Nepali and Maithili, which presented challenges in translations and maternal understanding. Transcripts were also subjected to several layers of translation. Qualitative interviewers listened to audio recordings of the interviews that were primarily conducted in Maithili and directly translated these into Nepali, which may have resulted in a loss of emic terms. Although clarifications were sought from a native Nepali speaker during analysis of English transcripts, the author of this thesis is a non-native Nepali speaker, further limiting the conclusions drawn from this analysis.

6.3 Programmatic implications

The findings from this thesis establish the need for further validation studies for the LBW and preterm birth indicators across a variety of settings, especially in populations that may be excluded if they do not seek care from facilities. If the factors associated with LBW and preterm birth overlap with those associated with a decreased likelihood of delivering in a facility, then studies utilizing facility-based designs are vulnerable to selection bias. Facility-based studies may then be unable to validate these indicators in subpopulations that may be the most vulnerable with a higher prevalence of LBW and preterm birth. In addition, these results indicate that tracking LBW over time using survey data at the same time as proportion of women delivering in facilities is increasing should be viewed with caution.

These results may help to inform the implementation of global household survey programs, like the DHS and MICS, to add to potential limitations in the interpretation of this type of data, and to provide preliminary evidence for the application of new methods to adjust the LBW indicator. The quantitative results show that data collected from maternal reports in rural Nepal result in underestimates of LBW and preterm birth. When compared to descriptive statistics presented in the 2011 Nepal DHS, 30.5% (343/1126) of newborns in population sampled from the central *terai* subregion, which encompasses Sarlahi District, had a reported birthweight.[1] From these, only 5.8% were classified as LBW, lower than the national average of 12.1%.[1] Of all newborns sampled in this subregion, 9.2% were identified as very small or smaller than average.[1] From the 2014 Nepal MICS, 47.8% of newborns in the central *terai* subregion were weighed at birth.[2] Of all newborns, 9.8% were described as smaller than average and none were very small.[2] Using the Blanc-Wardlaw correction, the LBW estimate was adjusted to 23.8%, lower than the national estimate of 24.2%.[2–4] These summaries fail to capture the much larger burden of LBW of almost 30% in Sarlahi found in the population-based parent trial.

As a result, program managers may not recognize the LBW burden affecting their regions and may divert resources to other health areas instead. On a global scale, if these results are similar to other rural areas, global household surveys may be systematically underestimating the LBW burden. With such dramatic improvements made in neonatal survival in recent years, accurate measurements become all the more important as efforts across the global community are required to be more focused on the most vulnerable communities.

6.4 Future research and next steps

We plan to publish separately findings from our validation of neonatal care seeking, postnatal care and immediate newborn care practices in this rural Nepali setting. As outlined by the Improving Coverage Measurement for MNCH group,[2] it is important to continue the work to validate measures of both intervention coverage and population in need indicators to ensure programmatic decisions are made using high-quality data. Validating these measures using various study designs and across different population settings will help us acquire a better understanding of what type of events mothers can accurately report and the direction and magnitude of possible biases. Using the visual aids developed for this project in a quantitative follow-on study may help to inform whether maternal responses for birth size can be improved. Finally, applying the new methods to multiply impute and adjust the LBW indicator to additional datasets will provide further information about the validity of this approach.

6.5 References

1. Ministry of Health and Population (MOHP) [Nepal], New ERA, ICF International Inc. Nepal Demographic and Health Survey 2011. Kathmandu, Nepal: 2012.

2. Central Bureau of Statistics. Nepal Multiple Indicator Cluster Survey 2014, Key Findings. Kathmandu, Nepal: Central Bureau of Statistics and UNICEF Nepal; 2014.

3. Blanc AK, Wardlaw T. Monitoring low birth weight: An evaluation of international estimates and an updated estimation procedure. Bull World Health Organ 2005;83(3):178–85.

4. United Nations Children's Fund and World Health Organization. Low Birthweight: Country, regional and global estimates. 2004.

5. Improving Coverage Measurement in MNCH. 2017 [cited 2017 Apr 29]; Available from: http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/current-projects/improving-coverage-measurements-for-mnch/

Appendices

7.1 Focus group discussion guide

Introduction

Namaste! Thank you participating in this group discussion today. I'm interested in hearing your opinions and experiences related to your work in the CVS study. This group discussion will be recorded using an audio device for research purposes.

Section A: Reflect on overall CVS work

1. How was your overall experience working on the CVS study? Did you have any specific challenges? Can you describe them?

Section B: Who answered the questions

1. When you were doing the interviews, do other people help mothers answer the questions, and if so, how do you manage this situation?

   Follow-up questions to probe for details:

   a. If other people tried to help the mother answer the questions, why do you think they tried to help answer the questions?

   b. If other people tried to help the mother answer the questions, how did this influence how you filled out the questionnaire?

   c. Do you think it is better to ask only the mother or to ask other people for help in answering the questions, and why?

Section C: Review of the questionnaire

First, please review the questions you asked mothers in the CVS study on the blank form provided to you.

1. Based on your experience, which questions do you think mothers had the most difficulty answering? *[Instruction: Note taker lists question numbers from the CVS questionnaire down on a copy of the questionnaire that WVCs identify as problems]*

2. *[For each of these questions]* Why do you think mothers found this difficult to answer?

   Follow-up questions to probe for details:

   Probe on understanding:

a. How do you think the phrasing of the question affects mothers' ability to understand this question?

b. If they did not understand the original wording of the question, what do you think mothers thought you were asking them?

c. How much did you have to probe for these mothers to understand what you were asking?

d. Without the probing, how do you think these mothers would have answered this question?

Probe on remembering:

e. How do you think time since the event affects mothers' ability to answer this question?

3. *[For questions mothers did not understand]* How can this question be improved to help mothers better understand what is being asked?

   Follow-up questions to probe for details:

   a. What are some common terms or phrases used in Sarlahi when discussing … *[insert the relevant question topic]*?

   b. How can we reword or change the words in this question to help mothers understand?

   c. How do you think using props (like a baby doll) would affect mothers' understanding? [*Refer to question 5 on the CVS questionnaire when discussing this probe]*

   d. How do you think pictures or drawings would affect mothers' understanding *[Show some examples of drawings]*?

4. *[For events mothers did not remember] How can this question be improved to help mothers better remember something that happened a while ago?*

   a. What are some ways to help mothers better remember dates or length of time?

   b. How do you think event calendars would affect mothers' recall? *[Explain what is meant by an event calendar – could use the earthquake as an example – did this happen before or after the earthquake, or did this happen during monsoon, or in the winter]*

Section D: Questions that were frequently discordant

[*This section is only needed if WVCs did not already identify these as problematic in Section C.]*

Now, let us discuss additional questions that based on the data that mothers answered differently in the CVS study compared to the NOMS main study.

1. *[For each of these questions]* Why do you think mothers found this difficult to answer?

   Follow-up questions to probe for details:

   Probe on understanding:
   a. How do you think the phrasing of the question affects mothers' ability to understand this question?

   b. If they did not understand the original wording of the question, what do you think mothers thought you were asking them?

   c. How much did you have to probe for these mothers to understand what you were asking?

   d. Without the probing, how do you think these mothers would have answered this question?

   Probe on remembering:
   e. How do you think time since the event affects mothers' ability to answer this question?

2. *[For questions mothers did not understand]* How can this question be improved to help mothers better understand what is being asked?

   Follow-up questions to probe for details:
   a. What are some common terms or phrases used in Sarlahi when discussing … *[insert the relevant question topic]*?

   b. How can we reword or change the words in this question to help mothers understand?

   c. How do you think using props (like a baby doll) would affect mothers' understanding? *[Refer to question 5 on the CVS questionnaire when discussing this probe]*

   d. How do you think pictures or drawings would affect mothers' understanding *[Show some examples of drawings]*?

3. *[For events mothers did not remember] How can this question be improved to help mothers better remember something that happened a while ago?*

    a. What are some ways to help mothers better remember dates or length of time?

    b. How do you think event calendars would affect mothers' recall? *[Explain what is meant by an event calendar – could use the earthquake as an example – did this happen before or after the earthquake, or did this happen during monsoon, or in the winter]*

Section D: Conclusion

1. Do you have any other issues related to the CVS study that you would like to discuss?

[STOP RECORDING]

7.2 In-depth interview guide

Introduction

Thank you again for agreeing to this interview. You will assist us in finding ways to help mothers like you better remember events related to their child's delivery and health soon after birth. There are no right or wrong answers. We are interested in hearing your opinion on this topic. This interview will be recorded using an audio device for research purpose, but all of your information will be kept private. If at any point during the interview you do not feel comfortable answering a question or wish to stop recording, please let me know.

****

Section A: Newborn Health

1. Is delivery and newborn health a topic you feel you can openly discuss in an interview? Why or why not?

2. When a Balposan worker first visited you after your child's birth, did other people help you answer questions about what happened? Why or why not?

3. If you wanted someone to check on your newborn's health, what specifically would you want them to do?

   Follow-up questions to probe for details:

   a. We describe a "health check" for a newborn as the following: "for example, someone examining (NAME OF CHILD), checking the cord, or seeing if (NAME OF CHILD) is OK." Can you explain how you might describe a "health check" for a newborn with other examples like these?

4. Can you please tell us your opinion on getting your newborn's health checked in the first 7 days after birth even if your newborn is not sick?

   Follow-up questions to probe for details:

   a. To prevent health problems during a high-risk period in the first 7 days after birth, someone should check on your newborn's health to make sure everything is okay. Do you agree or disagree with this statement? Why?

Section B: Time

1. After a child is born and before the placenta comes out, how aware do you think mothers like you are of the care their newborn is receiving?

   Follow-up questions to probe for details:

    a. For example, do you think most mothers can tell us if their newborn was wiped or washed with water before the placenta comes out? Why or not?

2. *[Example to assess how familiar women are with discussing time in hours]* Now I am going to ask you a series of questions that may seem unrelated to your child's health, but will help me understand other women's responses.

    Yesterday, when did you start preparing breakfast?
    How long did it take you to prepare breakfast?
    How long after your ate breakfast did you eat dinner?

   Now I am going to ask you a series of questions about when you went into labor.

    When did you start going into labor?
    How long after you went into labor was your child born?

3. We interviewed some mothers 1 month after birth and other mothers 24 months after birth. How do you think time that has passed since birth affects their ability to remember what happened?

****
[Complete Section C and D for each question that has a discordant answer listed for this participant.]

Section C: Reflect on discordant questions

1. *[For each of these questions, ask the mother the question and allow her to answer.]*

2. *[For each of these questions]* Can you please describe how you understood this question?

   Follow-up questions to probe for details:
    a. If you had to explain this question to a friend, how would you describe what the question is asking?

3. *[If the mother did not understand, explain the question and ask her to answer again.]*

Section D: How to improve questions

2. *[For Q3, Q4, Q5a, Q5b, show the mother the baby doll]* If I ask about health checks and use this doll to show that I am checking the baby's temperature or checking the baby's cord, how does this affect your understanding of the question? What do you like about this baby doll and what do you not like?

3. *[For Q17, show the mother the pictures]* If I ask about the size of your baby, and I show these pictures, how does this affect your understanding of the question? What do you like about these pictures and what do you not like?

4. *[For Q3 and Q4*] If I remind you about the holidays and festivals that happened in the time between your child's birth and 2 months after the birth, how does this affect your memory of what happened?

Section E: Conclusion

Do you have any other issues that you would like to discuss?
[Thank the mother for participating.]

Bibliography

Adegboye, A. R A, and B. L. Heitmann. 2008. "Accuracy and Correlates of Maternal Recall of Birthweight and Gestational Age." *BJOG: An International Journal of Obstetrics and Gynaecology* 115 (7): 886–93.

Alexander, G R. 2007. *"Prematurity at Birth: Determinants, Consequences, and Geographic Variation." Preterm Birth: Causes, Consequences, and Prevention*. Edited by Richard E Behrman and Adrienne S Butler. *Preterm Birth: Causes, Consequences and Prevention*. Washington D.C.: Institute of Medicine of the National Academies.

Araújo, Cora LP, Carmen LC Dutra, and Pedro C Hallal. 2007. "Validity of Maternal Report on Birth Weight 11 Years after Delivery: The 1993 Pelotas Birth Cohort Study, Rio Grande Do Sul State, Brazil." *Cadernos de Saúde Pública* 23 (10): 2421–27.

Bat-Erdene, Uilst, Amy Metcalfe, Sheila W McDonald, Suzanne C Tough, AR Adegboye, B Heitmann, SL Buka, et al. 2013. "Validation of Canadian Mothers' Recall of Events in Labour and Delivery with Electronic Health Records." *BMC Pregnancy and Childbirth* 13 (Suppl 1): S3. doi:10.1186/1471-2393-13-S1-S3.

Beatty, Paulc, and Gordonb Willis. 2007. "Research Synthesis: The Practice of Cognitive Interviewing." *Public Opinion Quarterly* 71 (2): 287–311.

Blanc, Ann K., and Tessa Wardlaw. 2005. "Monitoring Low Birth Weight: An Evaluation of International Estimates and an Updated Estimation Procedure." *Bulletin of the World Health Organization* 83 (3): 178–85.

Blanc, Ann K, Claudia Diaz, Katharine J Mccarthy, and Karla Berdichevsky. 2016. "Measuring Progress in Maternal and Newborn Health Care in Mexico: Validating Indicators of Health System Contact and Quality of Care." *BMC Pregnancy and Childbirth* 16. BMC Pregnancy and Childbirth: 1–11.

Blanc, Ann K, Charlotte Warren, Katharine J Mccarthy, James Kimani, Charity Ndwiga, and Saumya Ramarao. 2016. "Assessing the Validity of Indicators of the Quality of Maternal and Newborn Health Care in Kenya." *Journal of Global Health* 6 (1): 1–13.

Blencowe, Hannah, Simon Cousens, Mikkel Z. Oestergaard, Doris Chou, Ann Beth Moller, Rajesh Narwal, Alma Adler, et al. 2012. "National, Regional, and Worldwide Estimates of Preterm Birth Rates in the Year 2010 with Time Trends since 1990 for Selected Countries: A Systematic Analysis and Implications." *The Lancet* 379 (9832). Elsevier Ltd: 2162–72.

Boeke, Caroline E., Constanza Marín, Henry Oliveros, Mercedes Mora-Plazas, Samantha

Agudelo-Cañas, and Eduardo Villamor. 2012. “Validity of Maternal Birthweight Recall among Colombian Children.” *Maternal and Child Health Journal* 16 (4): 753–59.

Boerma, J. T., K. I. Weinstein, S. O. Rutstein, and A. E. Sommerfelt. 1996. “Data on Birth Weight in Developing Countries: Can Surveys Help?” *Bulletin of the World Health Organization* 74 (2): 209–16.

Braun, Virginia, and Victoria Clarke. 2006. “Using Thematic Analysis in Psychology.” *Qualitative Research in Psychology* 3 (2): 77–101.

Bryce, Jennifer, Fred Arnold, Ann Blanc, Attila Hancioglu, Holly Newby, Jennifer Requejo, and Tessa Wardlaw. 2013. “Measuring Coverage in MNCH: New Findings, New Strategies, and Recommendations for Action.” *PLoS Medicine* 10 (5): 1-9.

CDC National Center for Health Statistics. 2017. “Birthweight and Gestation.” https://www.cdc.gov/nchs/fastats/birthweight.htm.

Central Bureau of Statistics/Government of Nepal. 2015. “Nepal Multiple Indicator Cluster Survey.” Kathmandu, Nepal.

Central Bureau Of Statistics/Government of Nepal. 2012. “National Population and Housing Census 2011.” Vol. 1. Kathmandu, Nepal.

Channon, Andrew A. R. 2011. “Can Mothers Judge the Size of Their Newborn? Assessing the Determinants of a Mother’s Perception of a Baby’s Size at Birth.” *Journal of Biosocial Science* 43 (5): 555–73.

Channon, Andrew A R, Sabu S. Padmadas, and John W. McDonald. 2011. “Measuring Birth Weight in Developing Countries: Does the Method of Reporting in Retrospective Surveys Matter?” *Maternal and Child Health Journal* 15 (1): 12–18.

Charnigo, Richard, Lorie W Chesnut, Tony Lobianco, and Russell S Kirby. 2010. “Thinking Outside the Curve, Part I: Modeling Birthweight Distribution.” *BMC Pregnancy and Childbirth* 10 (37): 1-12.

Chen, Lincoln C, Emdadul Huq, and Stan D Souza. 1981. “Sex Bias in the Family Allocation of Food and Health Care in Rural Bangladesh.” *Population and Development Review* 7 (1): 55–70.

Creswell, John W. 2007. *Qualitative Inquiry and Research Design*. 2nd ed. Washington D.C.: Sage Publications, Inc.

Eggleston, E, A O Tsui, and J Fortney. 2000. “Assessing Survey Measures of Infant Birth Weight and Birth Size in Ecuador.” *Journal of Biosocial Science* 32 (3): 373–82.

Fereday, Jennifer, and Eimear Muir-Cochrane. 2006. “Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development.” *International Journal of Qualitative Methods* 5 (1): 80–92.

Gage, Timothy B. 2002. “Birth-Weight-Specific Infant and Neonatal Mortality: Effects of Heterogeneity in the Birth Cohort.” *Human Biology* 74 (2): 165–84.

Gage, Timothy B, and Gene Therriault. 1998. “Variability of Birth-Weight Distributions by Sex and Ethnicity: Analysis Using Mixture Models.” *Human Biology* 70 (3): 517–34.

Giri, Jamal, Prabhat Ranjan Pokharel, Rajesh Gyawali, and Bhushan Bhattarai. 2017. “Translation and Validation of Modified Dental Anxiety Scale: The Nepali Version.” *International Scholarly Research Notices* 2017. Hindawi Publishing Corporation: 1–5.

Grove, John, Mariam Claeson, Jennifer Bryce, Agbessi Amouzou, Ties Boerma, Peter Waiswa, Cesar Victora, and Kirkland Group. 2015. “Maternal, Newborn, and Child Health and the Sustainable Development Goals a Call for Sustained and Improved Measurement.” *Lancet* 386: 1511–14.

Gupta, Monica Das, Zhenghua Jiang, Bohua Li, Zhenming Xie, Woojin Chung, and Hwa-ok Bae. 2003. “Why Is Son Preference so Persistent in East and South Asia? A Cross-Country Study of China, India and the Republic of Korea.” *Journal of Development Studies* 40 (2): 153–87.

Hakim, R B, J M Tielsch, and L C See. 1992. “Agreement between Maternal Interview- and Medical Record-Based Gestational Age.” *American Journal of Epidemiology* 136 (5): 566–73.

Hancioglu, Attila, and Fred Arnold. 2013. “Measuring Coverage in MNCH: Tracking Progress in Health for Women and Children Using DHS and MICS Household Surveys.” *PLoS Medicine* 10 (5).

Harkness, Janet. 1999. “In Pursuit of Quality: Issues for Cross-National Survey Research.” *International Journal of Social Research Methodology* 2 (2): 125–40.

Hoffman, Caroline S., Lynne C. Messer, Pauline Mendola, David A. Savitz, Amy H. Herring, and Katherine E. Hartmann. 2008. “Comparison of Gestational Age at Birth Based on Last Menstrual Period and Ultrasound during the First Trimester.” *Paediatric and Perinatal Epidemiology* 22 (6): 587–96.

Howson, Christopher P, Mary V Kinney, Lori McDougall, and Joy E Lawn. 2013. “Born Too Soon: Preterm Birth Matters.” *Reproductive Health* 10 Suppl 1: S1.

Huang, Kristin T L, Claudio Owino, Rachel C Vreeman, Mildred Hagembe, Festus Njuguna, R Matthew Strother, and Gregory P Gramelspacher. 2012. “Assessment of the Face Validity of Two Pain Scales in Kenya : A Validation Study Using Cognitive Interviewing.” *BMC Palliative Care* 11 (5): 1–9.

“Improving Coverage Measurement in MNCH.” 2017. http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/current-projects/improving-coverage-measurements-for-mnch/.

"Indicator and Monitoring Framework for the Global Strategy for Women's, Children's and Adolescents' Health (2016-2030)." 2015. *Every Woman Every Child Sustainable Development Goals*.

INTERGROWTH-21st. 2017. "The International Fetal and Newborn Growth Consortium for the 21st Century." https://intergrowth21.tghn.org/.

Islam, M. Mazharul. 2014. "Can Maternal Recalled Birth Size Be Used as a Proxy Measure of Birth Weight? An Evaluation Based on a Population Health Survey in Oman." *Maternal and Child Health Journal* 18 (6): 1462–70.

Jowell, Roger. 1998. "How Comparative Is Comparitive Research?" *American Behavioral Scientist* 42 (2): 1998.

Kerm, Philippe Van. 2003. "Adaptive Kernel Density Estimation." In *9th UK Stata Users Meeting, Royal Statistical Society*. London, UK.

Kramer, M S. 1987. "Determinants of Low Birth Weight: Methodological Assessment and Meta-Analysis." *Bulletin of the World Health Organization* 65 (5): 663–737.

Lee, A C, J Katz, H Blencowe, S Cousens, N Kozuki, J P Vogel, L Adair, et al. 2013. "National and Regional Estimates of Term and Preterm Babies Born Small for Gestational Age in 138 Low-Income and Middle-Income Countries in 2010." *The Lancet.Global Health* 1 (1):26-36.

Li, Chung Yi, Jung Nan Wei, Tsung Hsueh Lu, Lee Ming Chuang, and Fung Chang Sung. 2006. "Mothers Tended to Overreport Categorical Infant Birth Weight of Their Children." *Journal of Clinical Epidemiology* 59 (12): 1319–25.

Liu, Li, Mengying Li, Li Yang, Lirong Ju, Biqin Tan, Neff Walker, Jennifer Bryce, Harry Campbell, Robert E. Black, and Yan Guo. 2013. "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* 8 (5).

Liu, Li, Shefali Oza, Dan Hogan, Yue Chu, Jamie Perin, Jun Zhu, Joy E Lawn, Simon Cousens, Colin Mathers, and Robert E Black. 2016. "Global, Regional, and National Causes of under-5 Mortality in 2000–15: An Updated Systematic Analysis with Implications for the Sustainable Development Goals." *The Lancet* 388 (10063): 3027–35.

Lule, Swaib a., Emily L. Webb, Juliet Ndibazza, Margaret Nampijja, Lawrence Muhangi, Florence Akello, Muhammed Kakande, Robert Kizindo, and Alison M. Elliott. 2012. "Maternal Recall of Birthweight and Birth Size in Entebbe, Uganda." *Tropical Medicine and International Health* 17 (12): 1465–69.

Marchant, Tanya, Jennifer Bryce, Cesar Victora, Allisyn C Moran, Mariam Claeson, Jennifer Requejo, Agbessi Amouzou, Neff Walker, Ties Boerma, and John Grove. 2016. "Improved Measurement for Mothers, Newborns and Children in the Era of the Sustainable Development Goals." *Journal of Global Health* 6 (1): 10506.

March of Dimes, PMNCHS, Save the Children, and WHO. 2012. "Born Too Soon: The Global Action Report on Preterm Birth. Eds CP Howson, MV Kinney, JE Lawn." *World Health Organization*. Geneva.

Mbuagbaw, Lawrence, and Rosa Gofin. 2010. "Can Recall of Birth Size Be Used as a Measure of Birthweight in Cameroon?" *Paediatric and Perinatal Epidemiology* 24 (4): 383–89.

McCarthy, Katharine J, Ann K Blanc, Charlotte E Warren, James Kimani, Brian Mdawida, and Charity Ndwidga. 2016. "Can Surveys of Women Accurately Track Indicators of Maternal and Newborn Care? A Validity and Reliability Study in Kenya." *Journal of Global Health* 6 (2).

Ministry of Health and Population (MOHP) [Nepal], New ERA, and ICF International Inc. 2012. "Nepal Demographic and Health Survey 2011." Kathmandu, Nepal.

Ministry of Health Nepal, New ERA, and ICF. 2017. "Nepal Demographic and Health Survey 2016: Key Indicators." Kathmandu, Nepal.

Moran, Allisyn C., Kate Kerber, Deborah Sitrin, Tanya Guenther, Claudia S. Morrissey, Holly Newby, Joy Fishel, P. Stan Yoder, Zelee Hill, and Joy E. Lawn. 2013. "Measuring Coverage in MNCH: Indicators for Global Tracking of Newborn Care." *PLoS Medicine* 10 (5): 1–9.

Moreno, L, and N Goldman. 1990. "An Assessment of Survey Data on Birthweight." *Social Science & Medicine (1982)* 31 (4): 491–500.

Mullany, Luke C, Gary L Darmstadt, Subarna K Khatry, and James M Tielsch. 2005. "Traditional Massage of Newborns in Nepal: Implications for Trials of Improved Practice." *Journal of Tropical Pediatrics* 51 (2): 82–86.

O'Sullivan, J J, M S Pearce, and L Parker. 2000. "Parental Recall of Birth Weight: How Accurate Is It?" *Archives of Disease in Childhood* 82 (3): 202–3.

Poulsen, Gry, Jennifer J. Kurinczuk, Dieter Wolke, Elaine M. Boyle, David Field, Zarko Alfirevic, and Maria A. Quigley. 2011. "Accurate Reporting of Expected Delivery Date by Mothers 9 Months after Birth." *Journal of Clinical Epidemiology* 64 (12): 1444–50.

Risnes, Kari R., Lars J. Vatten, Jennifer L. Baker, Karen Jameson, Ulla Sovio, Eero Kajantie, Merete Osler, et al. 2011. "Birthweight and Mortality in Adulthood: A Systematic Review and Meta-Analysis." *International Journal of Epidemiology* 40 (3): 647–61.

Robles, Arodys, and Noreen Goldman. 1999. "Can Accurate Data on Birthweight Be Obtained from Health Interview Surveys?" *International Journal of Epidemiology* 28 (5): 925–31.

Rosenstock, Summer, Joanne Katz, Luke C. Mullany, Subarna K. Khatry, Steven C. LeClerq, Gary L. Darmstadt, and James M. Tielsch. 2015. "Sex Differences in

Morbidity and Care-Seeking during the Neonatal Period in Rural Southern Nepal." *Journal of Health, Population and Nutrition* 33 (1): 1–10.

Rutstein, Shea Oscar, and Guillermo Rojas. 2006. "Guide to DHS Statistics. Demographic and Health Surveys Methodology." Calverton, Maryland.

Salasibew, M M, G Dinsa, D Berhanu, S Filteau, and T Marchant. 2015. "Measurement of Delayed Bathing and Early Initiation of Breastfeeding: A Cross-Sectional Survey Exploring Experiences of Data Collectors in Ethiopia." *BMC Pediatrics* 15 (1): no pagination.

Salasibew, Mihretab Melesse, Suzanne Filteau, and Tanya Marchant. 2014. "Measurement of Breastfeeding Initiation: Ethiopian Mothers' Perception about Survey Questions Assessing Early Initiation of Breastfeeding." *International Breastfeeding Journal* 9 (13): 1-8.

Salihu, Hamisu M. n.d. "The Historic Origin of Prematurity/Low Birth Weight: Lessons from the Past." *American Academy of Pediatrics*.

Schafer, Joseph L., and John W. Graham. 2002. "Missing Data: Our View of the State of the Art." *Psychological Methods* 7 (2): 147–77.

Secretaria de Salud Dirreccion General de Informacion en Salud. 2016. "Databases of Live Birth Certificates." http://www.dgis.salud.gob.mx/contenidos/basesdedatos/std_nacimientos_gobmx.html.

Seidman, D S, P E Slater, P Ever-Hadani, and R Gale. 1987. "Accuracy of Mothers' Recall of Birthweight and Gestational Age." *British Journal of Obstetrics and Gynaecology* 94 (8): 731–35.

Sou, Sio Choi, Wei J. Chen, Wu Shiun Hsieh, and Suh Fang Jeng. 2006. "Severe Obstetric Complications and Birth Characteristics in Preterm or Term Delivery Were Accurately Recalled by Mothers." *Journal of Clinical Epidemiology* 59 (4): 429–35.

Stanton, Cynthia K., Barbara Rawlins, Mary Drake, Matias dos Anjos, David Cantor, Lidia Chongo, Leonardo Chavane, Maria da Luz Vaz, and Jim Ricca. 2013. "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* 8 (5): 1-11.

Subramanyam, Malavika A., Leland K. Ackerson, and S. V. Subramanian. 2010. "Patterning in Birthweight in India: Analysis of Maternal Recall and Health Card Data." *PLoS ONE* 5 (7): 1–9.

Thapa, Shyam, Mala Dhital, and Shailes Neupane. 2002. "Assessing the Quality of Survey Data on Adolescent Sexuality by Talking with the Field Staff." *Contributions to Nepalese Studies* 29 (1): 1–30.

The World Bank Group. 2017. "The World Bank: Population, Total: Nepal." http://data.worldbank.org/indicator/SP.POP.TOTL?locations=NP.

The World Health Organization (WHO). 1995. "Physical Status: The Use and Interpretation of Anthropometry. Report of a WHO Expert Committee. Technical Report Series No. 854." *WHO Technical Report Series*. doi:854.

The World Health Organization (WHO). 2013. "Postnatal Care of the Mother and Newborn 2013." Geneva, Switzerland.

Tomeo, C A, J W Rich-Edwards, K B Michels, C S Berkey, D J Hunter, A L Frazier, W C Willett, and S L Buka. 1999. "Reproducibility and Validity of Maternal Recall of Pregnancy-Related Events." *Epidemiology (Cambridge, Mass.)*.

Umbach, David M, and Allen J Wilcox. 1996. "A Technique for Measuring Epidemiologically Useful Features of Birthweight Distributions." *Statistics in Medicine* 15: 1333–48.

Unicef. 2014. *Committing to Child Survival : A Promise Renewed. Unicef Progress Report 2014*.

United Nations Children's Fund (UNICEF). 2013. "Improving Child Nutrition. The Achievable Imperative for Global Progress." New York, NY.

United Nations Children's Fund (UNICEF). 2016. *The State of the World's Children 2016: A Fair Chance For Every Child*. New York, NY: United Nations Children's Fund (UNICEF).

United Nations Children's Fund (UNICEF). 2017. "Multiple Indicator Cluster Surveys." http://mics.unicef.org/.

United Nations Children's Fund and World Health Organization. 2004. *Low Birthweight: Country, Regional and Global Estimates*. *Unicef*.

United States Agency for International Development (USAID). 2014. "The DHS Program." *The Demographic and Health Surveys*. http://www.dhsprogram.com/.

Vecchio, Thomas J. 1966. "Predictive Value of a Single Diagnostic Test in Unselected Populations." *The New England Journal of Medicine* 274 (21): 1171–73.

Villar, José, Edgardo Abalos, Guillermo Carroli, Daniel Giordano, Daniel Wojdyla, Gilda Piaggio, Liana Campodonico, et al. 2004. "Heterogeneity of Perinatal Outcomes in the Preterm Delivery Syndrome." *Obstetrics and Gynecology* 104 (1): 78–87. doi:10.1097/01.AOG.0000130837.57743.7b.

Villar, José, Leila Cheikh Ismail, Cesar G Victora, Eric O Ohuma, Enrico Bertino, Doug G Altman, Ann Lambert, et al. 2014. "International Standards for Newborn Weight, Length, and Head Circumference by Gestational Age and Sex: The Newborn Cross-Sectional Study of the INTERGROWTH-21st Project." *The Lancet* 384 (9946): 857–68.

West, K P, R P Pokhrel, J Katz, S C LeClerq, S K Khatry, S R Shrestha, E K Pradhan, J M Tielsch, M R Pandey, and a Sommer. 1991. “Efficacy of Vitamin A in Reducing Preschool Child Mortality in Nepal.” *Lancet* 338 (8759): 67–71.

WHO. 2014. “Global Nutrition Targets 2025: Low Birth Weight Policy Brief.” *World Health Organization*. Geneva, Switzerland.

Wilcox, a J. 2001. “On the Importance--and the Unimportance--of Birthweight.” *International Journal of Epidemiology* 30 (6): 1233–41.

Wilcox, Allen J, and Ian T Russell. 1983. “Birthweight and Perinatal Mortality: II. On Weight-Specific Mortality.” *International Journal of Epidemiology* 12 (3): 319–25.

Wilcox, Allen J, and Ian T Russellf. 1983. “Birthweight and Perinatal Mortality: I. On the Frequency Distribution of Birthweight.” *International Journal of Epidemiology* 12 (3): 314–18.

World Health Organization (WHO);, and United Nations Children’s Fund (UNICEF). 2014. “Every Newborn: An Action Plan to End Prevantable Deaths.”

Yoder, P. Stanley, Mike Rosato, Riad Mahmud, Alfredo Fort, Faslur Rahman, Avril Armstrong, and Sayed Rubayet. 2010. “Women’s Recall of Delivery and Neonatal Care in Bangladesh and Malawi: A Study of Terms, Concepts and Survey Questions.” *DHS Qualitative Research Studies No. 17*. Calverton, Maryland, USA.

Yoder, P Stanley, and Laura Nyblade. 2004. “Comprehension of Questions in the Tanzania AIDS Indicator Survey .” *DHS Qualitative Research Studies No. 10*. Calverton, Maryland, USA.

Zou, Guangyong. 2004. “A Modified Poisson Regression Approach to Prospective Studies with Binary Data.” *American Journal of Epidemiology* 159 (7): 702–6.

Curriculum Vitae

KAREN TING CHANG

615 N.Wolfe St. Room W5031
Baltimore, MD 21204
kchang42@jhu.edu
(626) 233-3201

EDUCATION

Doctor of Philosophy (PhD), Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, Expected July 2017

- Department of International Health, Global Disease Epidemiology and Control (GDEC)

Master of Health Science (MHS), Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, May 2010

- Department of International Health, Global Disease Epidemiology and Control (GDEC)
- Certificate in Vaccine Science and Policy

Bachelor of Arts (BA), Public Health, University of California Berkeley, Berkeley, CA, May 2006

RELEVANT WORK EXPERIENCE

Research Assistant, Johns Hopkins Bloomberg School of Public Health, Department of International Health, Global Disease Epidemiology and Control, Sarlahi District, Nepal & Baltimore, MD, January 2016 – Current

- Develop a research proposal to validate maternal recall of postnatal care for newborns as part of a doctoral thesis project
- Conduct trainings of more than 40 local data collectors in quantitative and qualitative research methods
- Coordinate primary data collection activities in Nepal, including seeking IRB approvals, designing data collection tools, supervising data collectors, communicating project status updates to faculty in Baltimore, MD
- Analyze quantitative and qualitative results and report findings to the study sponsor and in a final dissertation

Teaching Assistant/ Floating Section Leader, Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology, Baltimore, MD, January 2017 – May 2017

- Lead undergraduate-level students enrolled in a Fundamentals in Epidemiology course through practice exercises to reinforce lecture material in one hour-long discussion sessions
- Address students' questions about course material during office hours
- Grade exams and assignments

Tutor in Biostatistics, Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, September 2014 – December 2015, January 2017 – March 2017

- Tutor master- and doctoral-level graduate students enrolled in the biostatistics 620 course series at the Johns Hopkins Bloomberg School of Public Health

Research Assistant, Johns Hopkins Bloomberg School of Public Health, Department of International Health, Human Nutrition, Baltimore, MD, February 2014 – January 2015; May – August 2015; May 2016 – August 2016

- Assisted on a situational analysis of the condition of adolescent girls in Burkina Faso, Indonesia, Zambia, and Tanzania to inform the World Food Programme (WFP) on strategies to improve the health status of female adolescents
- Analyzed Demographic and Health Survey (DHS) data from the listed countries to understand nutrition, reproductive and sexual health and education trends over time and disaggregated by urban/rural and wealth quintile variables
- Conducted a literature review to gather recent available information on adolescent health in these regions

Independent Consultant, Baltimore, MD September 2015 – July 2016

- WHO June 2016 – July 2016: Served as a rapporteur for a dengue vaccine licensing meeting at the World Health Organization, Geneva, Switzerland
- WHO November 2015 – January 2016: Drafted a Japanese Encephalitis (JE) vaccine guidance document targeting public health leaders in countries planning to introduce JE vaccines on best practices to measure vaccine effectiveness and vaccine impact
- Save the Children December 2015 – January 2016: Calculated and updated newborn health indicators with newly published data for the Healthy Newborn Network, which provides a comprehensive spreadsheet of the most recently published newborn health related indicators assembled from a variety of data sources
- Helen Keller International October 2015 – November 2015: Provided statistical guidance and conducted a revised statistical approach for a post-hoc analysis in a multi-country study assessing food consumption among children less than two years of age and maternal exposure to promotion for infant formula and complementary foods
- UNICEF September 2015 – November 2015: Led the analysis of literature gathered for a rapid qualitative review examining the use of the term "demand" and related terms in the context of vaccines and immunization for the Global Vaccine Action Plan at the WHO

Summer Placement, Jhpiego, Dar es Salaam, Tanzania, June 2014 – August 2014

- Conducted evaluation of the cervical cancer component of the Mothers and Infants, Safe, Healthy and Alive (MAISHA) program carried out in the Morogoro and Iringa regions from June 2010 to May 2014
- Analyzed program data including de-identified, client-level data of diagnoses and treatment, training records and previous annual reports
- Completed site visits to health facilities providing cervical cancer screening and treatment services to understand patient flow

ASPPH/CDC Allan Rosenfield Global Health Fellow, Association of Schools and Programs of Public Health (ASPH) / Centers for Disease Control and Prevention (CDC), Dar es Salaam, Tanzania, September 2011 – September 2013

- Conducted monitoring and evaluation research activities in the laboratory and blood safety branches, such as evaluation of a point-of-care CD4 testing technology for HIV diagnostics
- Contributed significantly to protocol development, coordination of training and implementation activities for evaluation studies, data management and analysis
- Built capacity in executing monitoring and evaluation activities of local counterparts of implementing partners
- Contributed scientific writing for abstracts and manuscripts

Malaria and TB Fellow, Global Health Access Program (GHAP), Mae Sot, Thailand, June 2010 – July 2011

- Assisted community-based organizations serving internally-displaced populations in Karen State, Burma in expanding malaria and TB screening and treatment programs
- Facilitated the development and advancement of program management and analytical skills of local partners through daily one-on-one capacity-building activities
- Assisted local partners in the development and implementation of curriculum for 6-8 day bi-annual training workshops of 8-15 local health workers from Karen State in Burma for their malaria and TB programs
- Drafted new protocols to investigate the risk of anti-malarial resistance in local communities through monitoring parasite clearance time

Research Assistant, Armed Forces Research Institute of Medical Sciences (AFRIMS), Department of Enteric Diseases, August 2009 – December 2009

- Analyzed data and drafted manuscripts on studies looking to establish novel challenge models for future *Shigella* vaccine trials in Thai adults and in rhesus monkeys conducted in Thailand and epidemiologic investigations of Rotavirus in children in Cambodia
- Rotated through laboratory sections for enteric diseases including bacteriology and molecular lab sciences
- Completed visits to sites conducting research on dengue fever and avian/swine flu, malaria, and diarrhea in Kamphaeng Phet and Mae Sot, Thailand and Phnom Penh, Cambodia

Summer Intern, PATH, Enteric Vaccines Initiative (EVI), Washington DC, June 2009 – August 2009

- Conducted a literature review of recently published data on the global epidemiology of diarrhea and the etiology of diarrhea-causing pathogens, specifically in low and middle-income countries
- Assisted in generating vaccine demand estimations by various populations, including travelers from developed countries, military, and endemic low- and middle-income countries, for potential future Shigella and ETEC vaccines using proprietary software developed by PATH

Research Assistant, Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology, Cochrane Eye Group, Baltimore, MD, January 2009 – August 2009

- Abstracted data relevant to trial design, methods of statistical analysis, and study outcomes from vision science articles and abstracts
- Conducted manual searches of vision science journals and conference proceedings for reports of randomized and controlled clinical trials
- Created a reference manager database of reports found describing randomized and controlled clinical trials

Ophthalmic Image Grading Specialist, Doheny Eye Institute, Los Angeles, CA, October 2006 – July 2008

- Analyzed retinal images for 10 clinical trials of pharmaceutical and government sponsorships
- Restructured the DIRC's in-house image-analysis training program
- Mentored 7 medical fellows and residents in completing the in-house training program to become certified in image grading
- Supervised and trained 2 junior image grading specialists

- Co-authored image-analysis protocols defining disease features as analyzed using proprietary software programs

PEER-REVIEWED PUBLICATIONS

Bodhidatta L, Pitisuttithum P, Chamnanchanant S, **Chang KT,** Islam D, Bussaratid V, Venkatesan MM, Hale TL, Mason CJ. Establishment of a Shigella sonnei human challenge model in Thailand. Vaccine. 2012 Nov; 19;30(49):7040-5.

Sadda SR, Liakopoulos S, Keane PA, Ongchin SC, Msutta S, **Chang KT**, Walsh AC. Relationship between angiographic and optical coherence tomographic (OCT) parameters for quantifying choroidal neovascular lesions. Graefes Arch Clin Exp Ophthalmol. 2010 Feb; 248(2):175-84. .

Keane PA, **Chang K**T, Liakopoulos S, Jivrajka RV, Walsh AC, Sadda SR. Effect of Ranibizumab Retreatment Frequency on Neurosensory Retinal Volume in Neovascular AMD. Retina 2009; 29(5):592-600.

Keane PA, Liakopoulos S, Jivrajka RV, **Chang KT**, Alasil T, Walsh AC, Sadda SR. Evaluation of optical coherence tomography retinal thickness parameters for use in clinical trials for neovascular age-related macular degeneration. Invest Ophthalmol Vis Sci 2009; 50(7):3378-85.

Keane PA, Liakopoulos S, **Chang KT**, Wang M, Dustin L, Walsh AC, Sadda SR. Relationship between Optical Coherence Tomography Retinal Parameters and Visual Acuity in Neovascular Age-Related Macular Degeneration. Ophthalmology. 2008;115:2206-14.

Keane PA, Liakopoulos S, **Chang KT**, Heussen FM, Ongchin SC, Walsh AC, Sadda SR. Comparison of the optical coherence tomographic features of choroid neovascular membranes in pathological myopia versus age-related macular degeneration, using quantitative subanalysis. Br J Ophthalmol 2008; 92:1081-5.

Keane PA, Liakopoulos S, Ongchin SC, Heussen FM, Msutta S, **Chang KT**, Walsh AC, Sadda SR. Quantitative Subanalysis of Optical Coherence Tomography after treatment with Ranibizumab for Neovascular Age-Related Macular Degeneration. Invest Ophthalmol Vis Sci 2008; 49:3115-20.

NON-PEER-REVIEWED PUBLICATIONS

De Pee S, **Chang K**, Ruel-Bergeron J. Improving Nutrition Among Adolescent Girls: Ways to reach them. *Sight and Life* 2016; 30(2): 94-100.

ABSTRACTS

Pathak S, Drammeh B, Mpopo G, Haule D, Mahmoud M, **Chang K,** Sembuche S, Kutaga R, De A. Real Time Quality Control and Data Management for the National Blood Transfusion Needs Study, Tanzania 2013. 2014 AABB Annual Meeting: Abstract No. A4-030B.

Drammeh B, Pathak S, Mpopo G, Haule D, Mahmoud M, **Chang K**, Sembuche S, Kutaga R, De A. Blood Demand and Estimation of Unmet Transfusion Needs, Tanzania. 2014 African Society for Blood Transfusion Conference: oral presentation.

Apata, I, Drammeh B, Pathak S, Mpopo G, Haule D, Mahmoud M, **Chang K**, Sembuche S, Kutaga R, De A. Clinical Utilization of Blood in Tanzania, June–September 2013. 2014 African Society for Blood Transfusion Conference: oral presentation.

Pathak S, Drammeh B, Mpopo G, Haule D, Mahmoud M, **Chang K,** Sembuche S, Kutaga R, De A. Data Management Operations for the National Blood Transfusion Needs Study, Tanzania 2013. 2014 African Society for Blood Transfusion Conference: poster presentation.

Arnett N, **Chang K**, Schmitz M, Lemwayi R, Rwehumbiza P, Mwasekaga M, Kohatsu L, Bolu O, Mosha F, Birhanu S, Lu L, Nkengasong J, Westerman L. Healthcare workers' acceptance and performance of point-of-care CD4 testing in Dar es Salaam, Tanzania, 2011. 7th International AIDS Society Conference: Abstract no. TUPDD0103.

Kohatsu L, Bolu O, Mosha F, Smith M, Lemwayi R, Arnett N, **Chang K**, Rwehumbiza R, Birhanu S, Sabatier J, Mwasekaga M, Westerman L, Nkengasong J. Evaluation of the point-of-care Pima CD4 assay and specimen collection methods at PMTCT and CTC sites in Dar-es-Salaam, Tanzania, 2011. 19th International AIDS Conference: Abstract no. LBPE17.

Chang KT, ShweeOo E, Kwee T, Smith L, Richards A, Lee TJ. Expansion of a malaria control program in Eastern Burma's active-conflict area and the need for a community-based approach. American Public Health Association 2011 139: E-Abstract 249921.

Chang KT ,Keane PA, Joeres S, Heussen FM, Ongchin SC, Walsh AC, Sadda SR. Analysis of a Putative Tachyphylaxis in Response to Treatment of Neovascular Age-Related Macular Degeneration with Intravitreal Ranibizumab. Invest Ophthalmol Vis Sci 2008 49: E-Abstract 341.

Sadda SR, Keane PA, Joeres S, Heussen FM, Ongchin SC, **Chang KT** , Walsh AC, . Quantitative Comparison of the Morphology of Choroidal Neovascularization in Pathologic Myopia versus that in Age-Related Macular Degeneration, using Optical Coherence Tomography Subanalysis. Invest Ophthalmol Vis Sci 2008 49: E-Abstract 263.

Keane PA, Joeres S, Ongchin SC, Heussen FM, Msutta S, **Chang KT** , Walsh AC, Sadda SR, . Quantitative Subanalysis of Optical Coherence Tomography Features after treatment with Ranibizumab for Neovascular Age-Related Macular Degeneration Invest Ophthalmol Vis Sci 2008 49: E-Abstract 5574

Sadda SR, Joeres S, Updike PG, **Chang KT**, Walsh AC. Quantitative Comparison of Short-Term Anatomic Response Following Intravitreal Bevacizumab versus Ranibizumab Treatment for Neovascular Age-Related Macular Degeneration. annual meeting of Macula Society, London, England. June 2007

Sadda SR, Joeres S, Updike PG, **Chang KT**, Walsh AC. Quantitative Comparison of Optical Coherence Tomography Data Following Intravitreal Bevacizumab, Ranibizumab or Pegaptanib Treatment for Neovascular Age-Related Macular Degeneration. Retina Society, 2007

HONORS AND AWARDS

Procter & Gamble Fellowship 2014 – 2015