Applying Multiple Imputation in the Analysis of Low Birth Weight in Nepal

Usha Singh

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Research Methodology

Prince of Songkla University

2016

Thesis Title Applying Multiple Imputation in the Analysis of Low Birth Weight in Nepal

Author Miss Usha Singh

Major Program Research Methodology

Major Advisor:	**Examining Committee:**
………………………………	……………………………Chairperson
(Dr. Attachai Ueranantasun)	(Asst. Prof. Dr. Nittaya McNeil)
	………………………………
Co-advisor:	(Dr. Attachai Ueranantasun)
………………………………………..........	……………………………………………
(Asst. Prof. Dr. Metta Kuning)	(Asst. Prof. Dr. Metta Kuning)
	…………………………………………
	(Assoc. Prof. Dr. Halimah Binti Awang)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Science Degree in Research Methodology.

……………………………………

(Assoc. Prof. Dr. Teerapol Srichana)

Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement have been made of any assistance received.

....................................Signature

(Dr. Attachai Ueranantasun)

Major Advisor

....................................Signature

(Miss Usha Singh)

Candidate

I hereby certify that this work has not been accepted in any substance for any degree, and is not being currently submitted in candidature for any degree.

....................................Signature

(Miss Usha Singh)

Candidate

ชื่อวิทยานิพนธ์ การประยุกต์ใช้การประมาณค่าทดแทนพหุในการวิเคราะห์ข้อมูลทารกแรกเกิดที่มีน้ำหนักต่ำกว่าเกณฑ์ในประเทศเนปาล

ผู้เขียน Miss Usha Singh

สาขาวิชา วิธีวิทยาการวิจัย

ปีการศึกษา 2558

บทคัดย่อ

ปัญหาสำคัญในการศึกษาเรื่องน้ำหนักทารกแรกเกิดคือการที่ข้อมูลน้ำหนักของทารกมักสูญหายจากการที่ไม่ได้ถูกบันทึกไว้แต่การศึกษาที่ผ่านมาไม่ได้พิจารณาวิธีนำข้อมูลน้ำหนักที่สูญหายมาประกอบในการวิเคราะห์ข้อมูล ดังนั้นวัตถุประสงค์ของการศึกษาครั้งนี้คือ เพื่อวิเคราะห์ปัจจัยต่างๆ ที่มีความสัมพันธ์ต่อน้ำหนักทารกแรกเกิดที่ต่ำกว่าเกณฑ์โดยใช้การประมาณค่าทดแทนพหุในการคาดคะเนข้อมูลที่สูญหายไป การศึกษาครั้งนี้ใช้ข้อมูลทุติยภูมิของทารกจากการสำรวจประชากรและสุขภาพในประเทศเนปาล)Nepal Demographic and Health Survey: NDHS)ในปี ค .ศ.2011การคาดคะเนข้อมูลที่สูญหายในการศึกษาครั้งนี้ใช้การประมาณค่าทดแทนพหุจำนวน65 ครั้ง โดยพบว่าจำนวนดังกล่าวทำให้ค่าความคลาดเคลื่อนของข้อมูลที่คาดคะเนได้มีค่าต่ำที่สุดทำให้ข้อมูลที่คาดคะเนมาเพื่อทดแทนข้อมูลที่สูญหายมีจำนวนทั้งสิ้น 65 ชุดและทุกชุดถูกนำมารวมกันเมื่อนำข้อมูลที่คาดคะเนรวมกับข้อมูลเดิมพบว่าทารกที่มีน้ำหนักแรกเกิดต่ำกว่าเกณฑ์คิดเป็นร้อยละ เพิ่มขึ้น 15.4 เมื่อพิจารณาปัจจัย 11.5เมื่อเทียบกับข้อมูลตั้งต้นที่มีเด็กทารกแรกเกิดที่มีน้ำหนักต่ำกว่าเกณฑ์ร้อยละแต่ละตัว พบว่าร้อยละของเด็กทารกแรกคลอดที่มีน้ำหนักต่ำกว่าเกณฑ์ จากข้อมูลก่อนและหลังจากการประมาณค่าทดแทนพหุมีความแตกต่างกันยกเว้นในปัจจัยอายุของมารดาขณะคลอด เพศของ

ทารก เชื้อชาติและที่อยู่อาศัยที่มีค่าใกล้เคียงกันสำหรับการวิเคราะห์ข้อมูลที่ได้หลังจากการคาดคะเนข้อมูลสูญหาย สถิติที่ใช้ในการวิเคราะห์ได้แก่ การวิเคราะห์การถดถอยพหุลอจิสติก ซึ่งถูกนำมาใช้เพื่อหาปัจจัยที่มีความสัมพันธ์ต่อน้ำหนักทารกแรกเกิดที่ต่ำกว่าเกณฑ์ผลการวิเคราะห์พบว่ากลุ่มมารดาที่มีสิทธิตัดสินใจเกี่ยวกับปัญหาสุขภาพของตนเอง มีแนวโน้มที่จะคลอดบุตรที่มีน้ำหนักต่ำกว่าเกณฑ์น้อยกว่ากลุ่มของมารดาที่มีสิทธิตัดสินใจเกี่ยวกับปัญหาสุขภาพโดยขึ้นอยู่กับสามีและบุคคลอื่น กับกลุ่มมารดาที่ตัดสินใจปัญหาสุขภาพร่วมกับสามี ผลจากการศึกษาพบว่าร้อยละของทารกแรกเกิดที่มีน้ำหนักต่ำกว่าเกณฑ์จากข้อมูลเริ่มต้นที่ยังไม่มีการคาดคะเนข้อมูลสูญหายมีแนวโน้มที่จะมีค่าน้อยเกินกว่าความเป็นจริง นอกจากนี้การศึกษายังพบว่าการวิเคราะห์ข้อมูลที่มีและไม่มีการประมาณค่าของข้อมูลที่สูญหายให้ผลที่แตกต่างกัน ดังนั้นการศึกษาเกี่ยวกับน้ำหนักทารกแรกเกิดที่ต่ำกว่าเกณฑ์ในควรจะมีการจัดการกับข้อมูลสูญหายไม่ว่าจะเป็นข้อมูลน้ำหนักของทารกและข้อมูลที่เกี่ยวข้องอื่น ๆ ผลจากการศึกษาครั้งนี้ยังเสนอประเด็นที่ควรมีการสนับสนุนให้เกิดขึ้นคือการส่งเสริมและสร้างความเข้มแข็งให้ผู้หญิงสามารถตัดสินใจด้วยตนเองได้ในเรื่องเกี่ยวกับสุขภาพของตนและควรมีการส่งเสริมและสร้างความเข้าใจในการใช้บริการด้านสุขภาพที่ถูกต้องให้กับมารดาและสามี รวมถึงครอบครัวอีกด้วยซึ่งจะส่วนช่วยลดความเสี่ยงของมารดาที่จะคลอดทารกที่มีน้ำหนักต่ำกว่าเกณฑ์ได้

Thesis Title	Applying Multiple Imputation in the Analysis of Low Birth Weight in Nepal
Author	Miss Usha Singh
Major Program	Research Methodology
Academic Year	2015

ABSTRACT

Studies conducted on birth weight have acknowledged missing data on birth weight, but these missing data are not included in the analysis. Furthermore, other existing missing data presented on determinants of birth weight are not addressed. Thus, this study is aimed to identify determinants that are associated with low birth weight (LBW) using multiple imputation to handle missing data on birth weight and its determinants. The child dataset from Nepal Demographic and Health Survey (NDHS), 2011 was utilized for this study. Multiple imputation carried out for 65 times was proven to be the lowest standard error than any other number of times of imputation. Therefore, each missing value was imputed for 65 times. The data showed prevalence of LBW as 11.5 percent, but the prevalence of LBW after imputation was 15.4 percent. The prevalences of LBW obtained from observed birth weight data set before and after imputation were different in each variable except for factors such as mother's age at child's birth, gender of child, ethnicity and residence. Multiple logistic regression was applied to find out the factors associated with LBW. Women with highest autonomy on their own health compared to those with involvement of husband or others, and with husband and women together were less likely to give birth to LBW infants. The findings of this study suggested that obtaining the

prevalence of LBW from only the sample of measured birth weight results in under estimation. In addition, assuming missing values as non missing provided different results from imputed. Therefore, it is suggested for future researchers conducting studies on LBW with DHS data from developing countries that missing data on birth weight and its determinants should be handled. The findings also suggested that there is a need for implementing programs that focus on promoting and strengthening women autonomy, as well as educational interventions not only to the couple but also to their in-laws regarding the importance of utilization of health services to help reduce the risk of birth to LBW infants among mothers.

Acknowledgements

I would like to express my sincere gratitude and deepest appreciation to my advisors Dr. Attachai Ueranantasun and Assist. Prof. Dr. Metta Kunning for their support and guidance throughout completion of this study. I am also grateful to Prof. Dr. Don McNeil, Assist. Prof. Dr. Phattrawan Tongkumchum, Assist. Prof. Dr. Apiradee Lim and Assist. Prof. Dr. Nittaya McNeil.

I am particularly thankful to Thailand's Education Hub for ASEAN Countries (TEH-AC) for providing financial support. I am also thankful to DHS measure for granting permission to conduct this study.

I want to express my profound appreciation to my colleagues in the Research Methodology program and the Department of Mathematics and Computer Science especially, Miss Jonu Pakhrin, Mr. Benjamin Owusu and Miss Nitinun Pongsiri.

Finally, I am thankful to my parents and rest of my family members for all the spiritual, emotional and financial support.

Usha Singh

Contents

List of Tables

List of Figures

List of Acronyms

UNICEF	The United Nations Children's Emergency Fund
WHO	World Health Organization
LBW	Low Birth Weight
NDHS	Nepal Demographic and Health Survey
DHS	Demographic and Health Survey
BMI	Body Mass Index
ANC	Antenatal Care
VDC	Village Development Committee
EA	Enumeration Area
MCAR	Missing Completely At Random
MAR	Missing at Random
MNAR	Missing Not at Random
JM	Joint Modeling
MICE	Multiple Imputation by Chained Equations
MID	Multiple Imputation then Deletion
FCS	Fully Conditional Specification
GLM	Generalized Linear Model

Chapter 1

Introduction

1.1 Background and rationale

Weight at birth determines baby's survival during their first 28 days and first year of life (Wilcox, 2001). Babies whose birth weight is low are 20 times more tended to die compared to heavier babies (UNICEF and WHO, 2004). Hence, birth weight is categorized into normal birth weight (babies weighing greater or equal to 2,500 grams) and low birth weight (babies weighing less than 2,500 grams) (World Health Organization, 2004). Low birth weight (LBW) infants are at a risk of developing diseases like diarrhea, respiratory infection (Lira *et al*., 1996), type II diabetes mellitus, cardiovascular diseases (Balci *et al*., 2010), renal failure (Fan *et al*., 2006) and mental disorder (Pedersen *et al*., 2013) during their early or late stage of life. Ultimately, due to health problems faced by LBW infants during their early or later stage of life, it is likely to increase significant cost to the society too (Almond *et al*., 2005).

Globally, 16 percent of all babies are born with LBW, accounting 95.6 percent births from the developing countries. The incidence of LBW has been constant in the past two decades in both Sub Saharan Africa and Asia where one out of four infants are born with LBW in South Asia (UNICEF, 2014). In Nepal, prevalence of low birth weight is 11.5 percent (MOHP and New ERA, 2011).

Several studies that are conducted on low birth weight by using demographic and health survey (DHS) data have found that factors like sex of the child

(Khanal *et al*., 2014; Sreeramareddy *et al*., 2011), birth order, hemoglobin level and cooking fuel (Sreeramareddy *et al.*, 2011), mother's education (Muula *et al*., 2011), antenatal visit during pregnancy, iron and folic acid supplementation during pregnancy (Khanal *et al*., 2014; Nisar & Dibley, 2014), geographical region (Khanal *et al*., 2014), residence and wealth index (Kayode *et al.*, 2014) are associated with LBW.

However, in developing countries, estimating percentage of LBW only from hospitals are biased, because only few deliveries take place at hospital, and mothers who deliver at hospital belong to high class are less likely to give LBW infants (Robles and Goldman, 1999). As a substitute to hospital based data, household survey data begin to collect information on infants born outside health facility (Blanc and Wardlaw, 2005). Since, a birth weight of infants is not measured at the time of birth, most of mothers are unable to provide numeric birth weight of their infants (Boerma *et al*., 1996). As an alternative to birth weight, it is suggested that mother's opinion on baby's birth size can be used (Boerma *et al*., 1996; Channon, 2011).

Birth weight is considered the main indicator of neonatal and infant health (Almond *et al*., 2002). In Nepal neonatal mortality rate is 33 per 100 births (MOHP and New ERA, 2011), accurate approximation of prevalence of LBW and its factors is necessary for intervening programs in reduction of infant and neonatal mortality.

MOHP and New ERA (2011) has reported that, in Nepal more than 50 percent of child's birth weight is not measured at the time of delivery because of home delivery. From the same survey report the prevalence of LBW is 11.5 percent, which is calculated only from the subsample of measured birth weight. Studies that are conducted in Nepal have identified the determinants that are associated with LBW only from the subset of measured birth weight (Khanal *et al*., 2014) or determinants that contribute small size at birth (Khanal *et al*., 2014). Besides missing values on the birth weight, missing values are also present on determinants of birth weight, but are not handled in most studies and the results obtained from these studies may be misrepresented. Estimating prevalence of LBW and identifying factors associated with it from the measured birth weight may not be justified (Robles and Goldman, 1999). Thus, the aim of this study is to identify factors associated with LBW using multiple imputations to handle missing data in both outcome and determinants.

1.2 Objectives

The objectives of this study are as follows:

a. To handle missing data available on both birth weight and its determinants using multiple imputation

b. To identify factors associated with low birth weight after imputation

1.3 Literature review

1.3.1 Low birth weight

At birth, baby's weight is considered the main factor that is directly associated to the health and nutrition of the mother (Wilcox, 2001). Birth weight is a single determinant that has direct impact on neonatal mortality. It is also considered to be a major determinant of infant and under five mortality and morbidity (Kramer, 1987). WHO (2004) had defined LBW as an infant's weighing less than 2,500 grams regardless of developmental age in mother's womb. Although neonatal mortality had declined gradually over past decades, the risk of dying low birth infants was 20 times higher than weighty babies (UNICEF and WHO, 2004). In addition, 60-80 percent infants born with LBW were likely to die within first 28 days (Lawn *et al*., 2005).

1.3.2 Institutional delivery in Nepal

In Nepal, 65 percent of the women gave birth at home (MOHP and New ERA, 2011). The government of Nepal had implemented number of interventions in order to increase the rate of hospital delivery and access to 24 hours emergency obstetric care through different incentive programs and establishment of birth centers in each district (Sreeramareddy *et al.*, 2011). In 2009, a national free delivery policy known as *Aama* program was introduced by government of Nepal. Under this program, mothers received cash rewards for attending at least four complete antenatal care (ANC) visits, total reimbursement of cost of transportation to a health facility for delivery and free delivery service (MOHP, 2011).

1.3.3 Data on low birth weight

In many developing countries, it was difficult to identify prevalence of LBW because most of the deliveries took place at home and birth weights were not measured (Blanc and Wardlaw, 2005). So, study conducted using hospital records to estimate LBW may not represent nationally. As an alternative, demographic and health survey (DHS) began to collect information on birth weight based on both birth certificate or mother's recall (Blanc and Wardlaw, 2005). However, this survey data became limited because most of the infants weights were not measured at the time of birth.

Boerma *et al*. (1996) used mother's recall on size at birth (i.e. very large, larger than average, average, smaller than average and very small) as an alternative to the birth weight for estimating LBW.

The study conducted to explore accuracy of birth weight information from six DHS data revealed that estimating low birth only from the measured birth weight sample from the surveys of developing countries may depict an exaggerative good picture of infant's and mother's health condition (Robles and Goldman, 1999). The findings of this study were supported by a similar study which analyzed 62 DHS of 42 developing countries. It was illustrated that mothers who gave birth at hospital were educated and lived in urban areas (Blanc and Wardlaw, 2005).

1.3.4 Heaping

Heaping or clustering of a data is an aggregation of large numbers of numeric values ending in 0 or 5 (Blanc and Wardlaw, 2005). Heaping occurs in birth weight data when health personnel or mother reports infant's weight on figures which are multiplication of 500 grams (Boerma *et al*., 1996). High level of heaping on birth weight data indicates poor mother's recall (Robles and Goldman, 1999). The incidence of LBW was under reported in developing countries, which was due to high degree of clustering at 2,500 grams and excluding all infants weighing 2,500 grams as LBW. In order to estimate the prevalence of LBW by lowering considerably bias, a study conducted on monitoring low birth weight, classified 25 percent of infants as LBW of those infants whose weights were reported exactly 2,500 grams (Blanc and Wardlaw, 2005). However, a similar study was conducted by analyzing 15 DHS surveys data, assuming half of the reported birth weight as 2,500 grams as LBW (Boerma *et al*., 1996).

1.3.5 Determinants

The underlying factors of LBW are not thoroughly studied, although some researchers have tried to identify them. Determinants that are found significant with LBW are mother's age at child's birth, mother's education, ANC visit and consumption of iron tablets during pregnancy, BMI, parity, smoking, gender of child, economic status, cooking fuel and residence.

Mother's age at child's birth

The study conducted in US, showed U shaped relation between the mother's age and the LBW among whites. The results showed that mothers with the age below 15 and above 40 were more inclined to give birth to LBW babies than mothers between age of 25-29 years (Reichman and Pagnini, 1997). However, a systematic review revealed an evidence of a dose-response relationship between mother's age at birth and LBW which means the degree of association decreased with the increase in mother's age (Gibbs *et al*., 2012). King (2003) explained how teenage mothers gave birth to LBW infants. The findings showed that in growing teens there was an increase of Leptin which decreased fat breakdown and increased mother's use of glucose for energy as a result lesser amount of glucose was supplied to the fetus. Similarly, during young age internal organs of reproductive system did not fully mature as a result teenagers were likely to give LBW infants (Gibbs *et al.,* 2012).

Mother's education

Several researchers found an inverse relationship between mother's education and LBW. A study on LBW found that the odds of delivering LBW was greater in those women who had no formal education compared to those who had at least secondary level of education (Muula *et al.*, 2011). Another study in Nepal also showed that infants with LBW were significantly higher among illiterate mothers than literate ones (Mondal.B, 2000). A meta-analysis study illustrated that, mothers with higher education had 33 percent constructive effect against LBW as compared to medium and low level education (Silvestrin *et al*., 2013).

Maternal body mass index (BMI)

Many studies revealed strong association between mother's BMI and birth weight of the newborns. A study conducted by Doherty *et al*. (2006) found that underweight mothers were considered risk factor for inhibiting fetal growth. This result was supported by a study conducted in Germany revealed that mothers with low BMI were more likely to give birth to preterm and LBW infants compared to mother with normal BMI (Kalk *et al*., 2009). A study conducted in Nepal, also found that low BMI increased the risk of LBW (Bondevik *et al.*,2001).

Parity

Several studies found that nulliparous and primi mothers were at a risk of giving birth to LBW infants. A meta-analysis study showed that nulliparous women aged less than 18 years were more tended to give birth to LBW infants compared to women who have 1-2 children with age between 18-34 years (Badshah *et al*., 2008). Similarly, a study conducted by Muula *et al.* (2011) found that mothers who had given birth before were less inclined to give birth to LBW infants.

Smoking

Smoking during pregnancy was considered a strong dose-dependent determinant for LBW and preterm birth (Kramer, 1987; Wilcox, 2001). A study conducted by Chiolero *et al*. (2005), illustrated that mothers who smoke were 2.7 times more inclined to give birth to LBW infants compared to non smokers.

Women's decision for utilization of health services

Women's participation for making decision on utilization of health of services was considered an indicator of women's autonomy. A study conducted in Bangladesh revealed that women with lowest decision making autonomy were more likely to give birth to LBW (Sharma and Kader, 2013). Furthermore, study conducted in Nepal showed that husbands had great control on the use of maternal health services among teenagers and young adults (Upadhyay *et al*., 2014).

Cooking fuel

A study conducted in India found that mothers using high pollution fuel were more inclined to have LBW babies than mothers using non pollution cooking fuel (Sreeramareddy *et al*., 2011). However, studies conducted in Nepal found that cooking fuel as a non risk factor for LBW (Khanal *et al*., 2014; Khanal *et al*., 2014).

Residence

A study to identify the risk factors for LBW in sub Saharan Africa reported that chance of LBW among mother living in rural areas was 43 percent due to inadequate health facilities, lack of job opportunities and infrastructure development (Kayode *et al*., 2014). In contrast to this finding, study from Brazil reported that prevalence of

LBW was higher in city dwellers than rural dwellers which was due to incomplete registration of live births and registration of live births as stillbirths in rural areas (Silva *et al*., 2005). However, another study in USA showed that prevalence of LBW was higher in crowded urban and remote rural areas (Kent *et al*., 2013).

Economic status

It was found that economic disadvantage was associated with low birth weight. For example, a study conducted in Ghana illustrated that women living in low economic status had the risk of having LBW infants twice than women living in wealthier communities (Kayode *et al*., 2014). The finding was consistent with a previous study conducted in Malawi which found that risk of delivering LBW infant was lower for mother having high economic status (Muula *et al*., 2011).

Gender of infant

Regarding gender of infant, it was revealed that female infants were born with LBW than male infants. A study conducted in India illustrated that incidence of LBW was greater among female babies compared to male babies (Mondal, 1998). Likewise, a study conducted in Nepal also reported that risk of being small in size was less among male infant than female infant (Khanal *et al*., 2014).

Number of antenatal care (ANC) visit during pregnancy

There have been studies showing that utilization of medical services during pregnancy, lessen poor birth outcomes. A study on small size at birth in Nepal reported that the risk of giving birth to small size babies was higher to those mothers who had no antenatal visits than those who had at least four or more antenatal visits

(Khanal *et al*., 2014). The findings of this study was consistent by a similar study conducted in Nepal found that odds of having LBW was increased by double among mothers who did not attend antenatal visits (Khanal *et al*., 2014). Likewise, a study conducted in India had identified significant association between antenatal care and LBW (Dharmalingam *et al*., 2010).

Iron consumptions during pregnancy

It was known from the several studies that hemoglobin level is decreased during pregnancy period. In order to prevent from iron deficiency anemia and promote healthy pregnancy outcomes, iron tablet was supplemented to the pregnant mother from the third trimester. A study conducted in Pakistan showed that, the chance of having smaller infants was considerably reduced by 18 percent for mothers who consumed iron folic tablets than those who did not consume (Nisar and Dibley, 2014). In the same way, a study conducted in Nepal also reported that mothers consuming iron tablet during pregnancy were less tended to have LBW infants compared to mothers not consuming iron tablets during their pregnancy (Khanal *et al*., 2014).

1.3.6 Statistical method

In most of the studies birth weight was treated as binary outcome i.e. birth weight greater or equal to 2,500 grams and birth weight less than 2,500 grams and used multiple logistic regression model to estimate the prevalence and factors associated with LBW (Khanal *et al*., 2014; Nisar and Dibley, 2014). However, a community based study conducted in China used multiple linear regression (continuous outcome) for identifying whether air pollution commonly had an effect on all births, which was calculated by decreased birth weight and multiple logistic regression (binary outcome) for identifying the impact of air pollution on LBW (Wang *et al*., 1997).

1.3.7 Missing data

Despite of designing study carefully, missing data still occur in most of the studies (van der Heijden *et al*., 2006). For a survey study, when a subject in a sample did not answer any of the questions, such type of missing is called as unit non response whereas when a sampled subject partially answers questions, such type of missing is called as item non response (Bethlehem, 2009). To handle above mentioned non-response data, usually two approaches are used namely imputation and adjustment weighting (Lumley, 2011). Imputation is used for management of item non-response data and weighting is used for the management of unit non response data (Schafer and Graham, 2002).

Mechanism of missing data

Mechanism of missing data is categorized into three; missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) (Rubin, 1987). For the data that are MCAR and MAR, the missing mechanism is also referred to as ignorable, whereas for data that is MNAR, the missing mechanism is referred to as non-ignorable (Rubin, 1976; van Buuren, 2012).

Missing completely at random (MCAR)

Rubin (1976) defined missing mechanism as MCAR, when the probability of missingness is independent of both observed and missing data. For example, missing data are MCAR when survey papers are lost accidentally.

To understand MCAR with notation, assuming that Y_{com} as complete data set, which is further divided into: Observed data set (Y_{obs}) and missing data set (Y_{mis}). Similarly, consider, I as matrix of missing values with the same dimension as data matrix (Y). I can be 0 or 1 depends upon Y. If Y is missing (Y_{mis}) or observed (Y_{obs}), it is coded as 0 and 1 respectively (Dong and Peng, 2013; Schafer and Graham, 2002). Thus, under this mechanism, the probability of missingness is independent of both Y_{obs} and Y_{mis} and can be written as (Schafer and Graham, 2002):

$$\mathrm{P}(I|Y_{\mathrm{com}}) = \mathrm{P}(I)$$

Missing at random (MAR)

The probability of missingness depends on Y_{obs} and not on $Y_{mis.}$ For example, people from different demographic backgrounds may decline to answer based on beliefs or traditions. It can be written as (Schafer and Graham, 2002):

$$\mathrm{P}(I|Y_{\mathrm{com}}) = \mathrm{P}(I|Y_{\mathrm{obs}})$$

Missing not at random (MNAR)

The probability of missingness depends on both missing (Y_{mis}) and observed (Y_{obs}) data. For example, high income people are less likely to report their income than that of people with average or low income (Dong and Peng, 2013)

1.3.8 Imputation

There are several studies about methods used for handling missing data in each type of missing mechanism (Bennett, 2001). The most common method is complete case analysis, usually referred as case deletion where missing values are deleted. Results from this method are inefficient, but unbiased, when the missing data hold MCAR assumption. However, when data are not MCAR, results from this method are inefficient and biased (Little and Rubin, 2002; Schafer and Graham, 2002).

If the missing data are not MCAR, imputation methods are implemented to handle missing data. Imputation is to predict and substitute an appropriate value for a missing value in a data set (De Waal *et al*., 2011). Methods like mean substitution, last observation carried forward, hot deck imputation, cold deck imputation and regression imputation come under single imputation in which missing values are substituted by appropriate values (Bethlehem, 2009; De Waal *et al*., 2011). Furthermore, in single imputation, values are imputed for one time; the uncertainties created by missing values are not accounted for. As a result there are small standard errors, p-values and narrow confidence interval (Little & Rubin, 2002; Sterne *et al*., 2009). In multiple imputation, unlike single imputation, missing values are imputed for more than one time and the uncertainties created by missing values are incorporated resulting into large standard errors and wider confidence intervals (Rubin, 1987). Although, multiple imputation results in larger standard error, it gives unbiased result (Schafer and Graham, 2002).

van der Heijden *et al.* (2006) compared several methods of imputations such as case deletion, missing indicator method, single imputation and multiple imputation for handling missing data. After imputation of missing values from four different methods, multivariable logistic regression was fitted to identify factors associated with outcome. The predictors of the outcome based on the case deletion and missing indicator methods varied from the predictors of the outcome based on other models. The findings suggested the avoidance of using case deletion and missing indicator method although the data are MAR.

Little (1988) stated that if the missing values are presented in both determinants (X) and outcome (Y), then cases with missing Y can carry only small information which is not very useful for regressing. Therefore, under such condition Multiple Imputation then Deletion (MID) performs better than standard multiple imputation in which all missing values on X and outcome Y are imputed, and then deleting cases with imputed values on Y before analysis (von Hippel, 2007). However, standard multiple imputation performs better than MID when auxiliary variables are employed in an imputation model (Sullivan *et al.*, 2015).

1.3.9 Sample weight

In demographic and health surveys, samples were chosen with uneven probability to increase the number of cases. Thus, sample weight was required for any analysis using DHS data to make sure findings represent nationally (MOHP and New ERA, 2011).

Sampling weights are adjustment factors that are used to adjust for the selection difference probabilities. The difference may be due to study design or happenstance (MOHP and New ERA, 2011). Studies conducted by using DHS survey data used Complex Sample Analysis method (Khanal *et al.*, 2014), multilevel modelling (Sreeramareddy *et al*., 2011) and "Svy" commands in STATA (Nisar and Dibley, 2014) to account for sampling design and sample weights.

1.4 Scope of the study

The analyses focused on 5,240 children excluding 66 multiple births.

The determinants of this study were maternal factors, child factors and socio-demographic factors. Data were obtained from Nepal Demographic and Health

Survey (2011). Multiple imputation was used to handle missing data. Simple and multiple survey logistic regressions were carried out to estimate prevalence and factors associated with LBW.

1.5 Conceptual framework of the study

The conceptual framework of the study is shown in Figure 1.1.

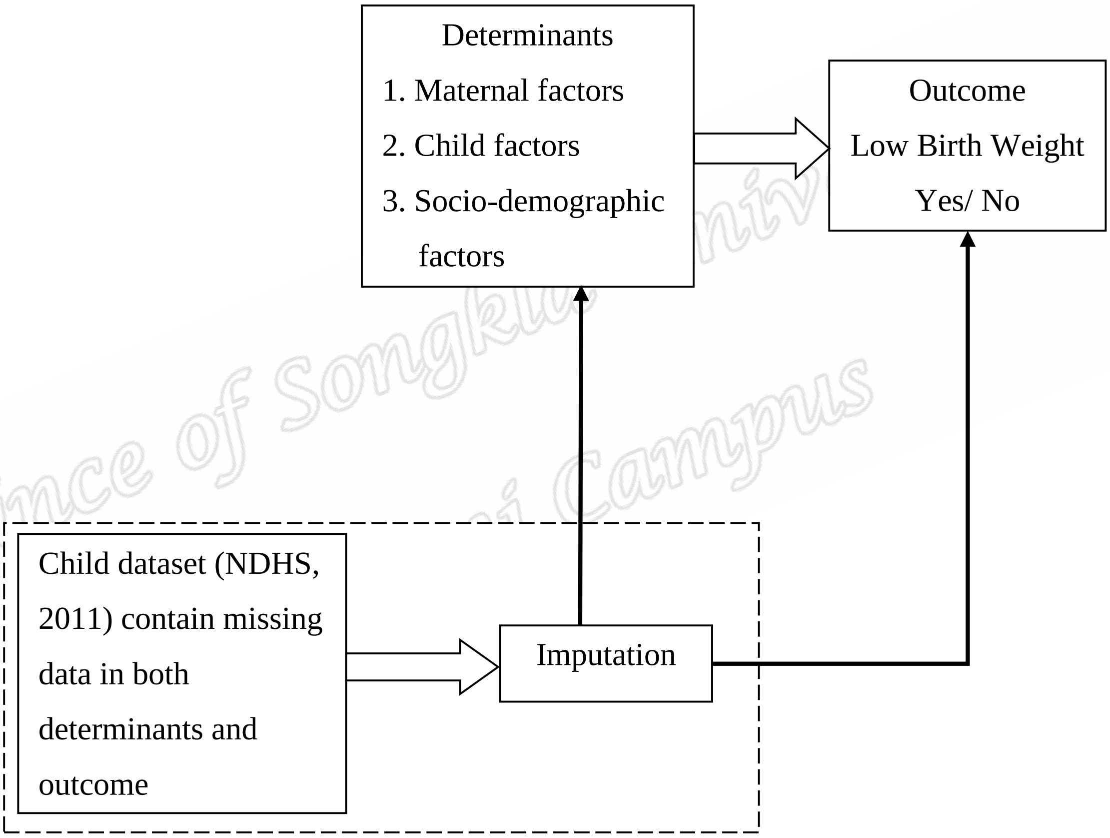

Figure 1.1 Conceptual framework of the study

Figure 1.1 shows the conceptual framework of the study. The child data set contain missing data in both determinants and outcome. Therefore, imputation is used to handle missing data. After handling missing data, survey logistic regression is applied to identify factors associated with LBW.

1.6 Flow chart of the study

The flow chart of this study is shown in Figure 1.2.

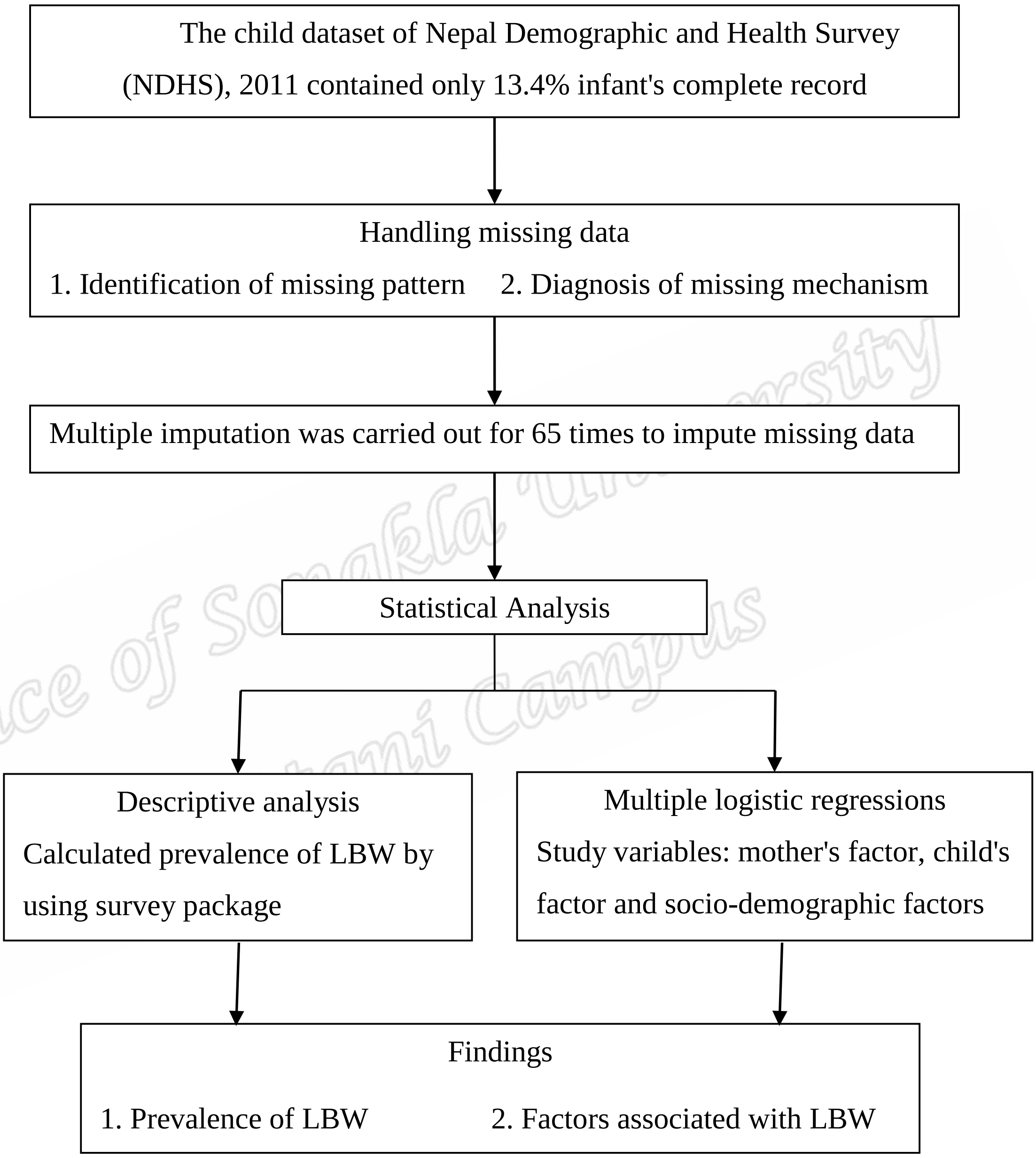

Figure 1.2 Flow chart of the study

Figure 1.2 presents flow chart of this study. The child data set contain only 13.4% of infant's complete record. Hence, multiple imputation is applied to handle missing data. However, before handling missing data, frequency, pattern and mechanism of missing data are identified. After imputation, descriptive statistics and multiple logistic regression is applied to identify prevalence and factors associated with LBW.

1.7 Organization of thesis

This thesis contains four chapters. The organization of these chapters is described below:

Chapter 1 describes background objectives, literature review, conceptual framework and scope of this study. Chapter 2 provides description of all methodologies used in this study. Chapter 3 presents prevalence of LBW in each subgroup and describes results from survey logistic regression model. Chapter 4 concludes and discusses the research findings.

Chapter 2

Methodology

This chapter consists of a two parts: the first part describes about data and a source of data, and how the data is collected and study sample. The second part describes about method that is used to handle missing data and statistical methods.

2.1 Data from NDHS, 2011

The child dataset from Nepal Demographic and Health Survey (NDHS), 2011, was used in this study. The 2011 NDHS was fourth nationally representative household survey conducted in every five years. The survey was carried out by the support of Ministry of Health and Population and was executed by New Era. Three sets of questionnaires were employed in the survey; the household questionnaire, woman's questionnaire and man's questionnaire. These questionnaires allowed different units of analysis and they were eventually converted into datasets. The types of datasets presented in this survey are described below:

Household recode (HR)

This dataset contains a record for each household which includes household members list. The component of analysis in this dataset is a household. It includes information on household characteristics such as housing condition, availability of drinking and toilet facilities, type of cooking fuel used, etc.

Household member recode (PR)

This file contains a record for every household member and the component of analysis is a household member. It includes information of household members such as age, sex, education, occupation, income, height and weight.

Individual recode (IR)

It contains a record for every eligible woman and the unit of analysis is a woman. Information collected from the women's questionnaire and few variables from the household questionnaire are included in this data set.

Male recode (MR)

This file contains a record for every eligible man as termed by the household. It includes all the information gathered in the men's questionnaire along with few variables from the household.

Couple's recode (CR)

This file contains information on each couple and the component of analysis is the couple. Variables such as age, sex, education, occupation and types of family planning used by couples are included in this dataset.

Children's recode (KR)

This file has information on each child of interviewed women, born in the period between 2006- 2011. It includes the information related to the child's birth weight, immunization and health status. The unit of analysis in this dataset is a child born within last 5 years.

Births recode (BR)

This dataset includes information on each child ever born to interviewed women. It includes the birth and birth history of all women interviewed. The unit of analysis is the children ever born from eligible women.

2.2 Sampling

Nepal is divided into 75 districts; each district is divided into municipalities and village development committees (VDCs). The VDCs and municipalities are again divided into wards and sub-wards according to population size. Thus, an enumeration area (EA) is a ward in rural areas and sub-wards in urban areas. Based upon ecological zones and development regions; the country is divided into three and five respectively. The cross-section of these zones and regions yields 15 eco-development regions. However, in this survey, the mountain zone of Western, Mid-western and Far-western were merged to make 13 sub-regions, because the population in these zones is very few.

The 2011 NDHS sample was selected by using a stratified two-stage cluster design. Stratification was done by the cross-classification of urban and rural areas in 13 sub-regions resulting into 25 strata. At first, probability proportionate to size was used to select EAs from each stratum. A total of 289 EAs were selected; 194 from rural areas and 95 from urban areas. In the second stage, random sampling was used to select of 40 households from each rural areas and 35 households from each urban areas. In NDHS (2011), 11,085 households were expected in the sample, but only 10,826 households were interviewed. The process of sampling is described in Figure 2.1.

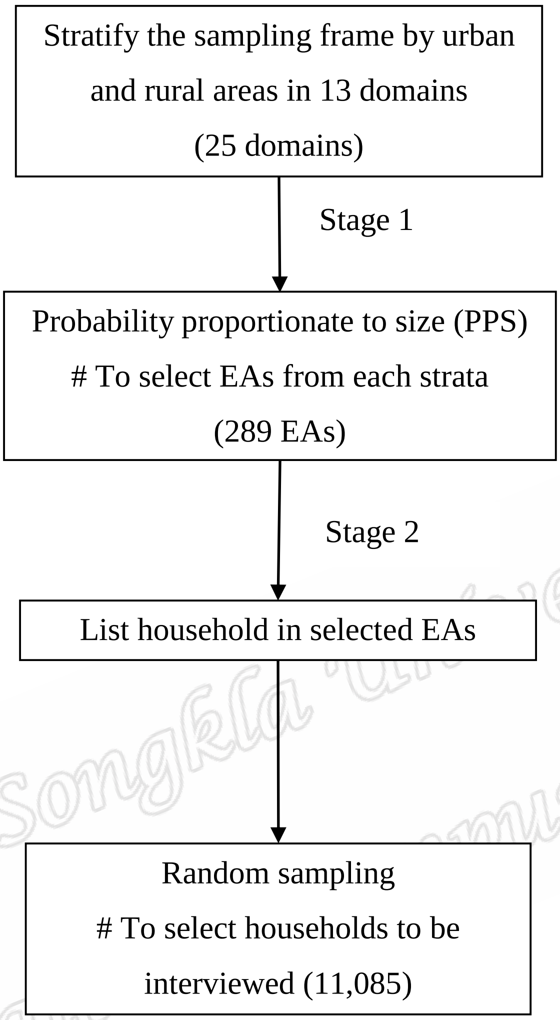

Figure 2.1 Process of sampling

However, the number of wards and sub wards in each of the 13 domains were not assigned proportionally to their population. Likewise, the greater number of the population lives in rural areas than urban areas. As a result, urban areas were oversampled and rural areas were under sampled in order to provide national estimates. However, due to oversampling and under sampling, the estimate for Nepal as a whole would be biased since the distributions of samples in the regions were dramatically different to that of actual distribution of Nepal. To restore national representativeness of samples, MOHP and New ERA, (2011) has recommended the use of sample weight before any statistical analysis.

2.3 Study sample

In this study, child dataset was used for its information on under-five year children. A total of 5,306 children were born during the period of 2006-2011. Children from multiple births tended to have lower birth weights than singletons. Furthermore, it was likely that correlations within sets of multiple births are higher than those between singletons (Khanal *et al*., 2014). Thus, 66 multiple births were excluded from this study and analyzed only 5,240 children.

2.4 Study variables

2.4.1 Dependent variable

Birth weight was considered binary outcome in this study and was classified into; birth weight greater or equal to 2,500 grams and birth weight less than 2,500 grams. Path diagram of the study variables are shown in Figure 2.2.

2.4.2 Independent variables

2.4.2.1 Maternal factors

Mother's age at child's birth was categorized as 15-19 years, 20-24, 25-29years and over 30 years.

Mother's education was grouped into no education, primary, secondary and higher education.

Parity was classified as first, second or third and fourth child or more.

Antenatal care (ANC) visit during pregnancy was classified as no visit, 1-3 visits and 4 or more visits.

Consumption of iron tablets during pregnancy was categorized as yes or no.

Mother's body mass index (BMI): BMI was grouped as underweight (< 18.5 km/m^2), normal (18.5 - 23.0 kg/m^2) and overweight (> 23.0 kg/m^2).

Women's decision for utilization of health services was grouped into women, women and husband together and husband or others.

Smoking tobacco was categorized as yes and no.

2.4.2.2 Child factors

Gender of child was categorized as male and female.

Birth interval was classified as no interval, less than 24 months and more or equal to 24 months.

2.4.2.3 Socio-economic factors

Wealth status was grouped into poor, middle and rich.

Ethnicity was classified as relatively advantaged, relatively disadvantaged (Janajati) and relatively disadvantaged (Dalit).

Cooking fuel was grouped into non polluting fuel and high polluting fuel.

Residence was categorized into rural and urban.

Ecological region was classified as Mountain, Terai and Hill.

Development region was categorized as Eastern region, Central region, Western region, Mid-western region and Far-western region.

2.4.3 Auxiliary variable

For mother's opinion on infant's size at birth, besides birth weight, NDHS 2011 asked a specific question to the mothers about the size of their babies at the time of birth. Based upon five categories namely very large, large, normal, small and very small, mothers had to recall their babies' size. This variable was used as an auxiliary variable while imputing birth weight in this study.

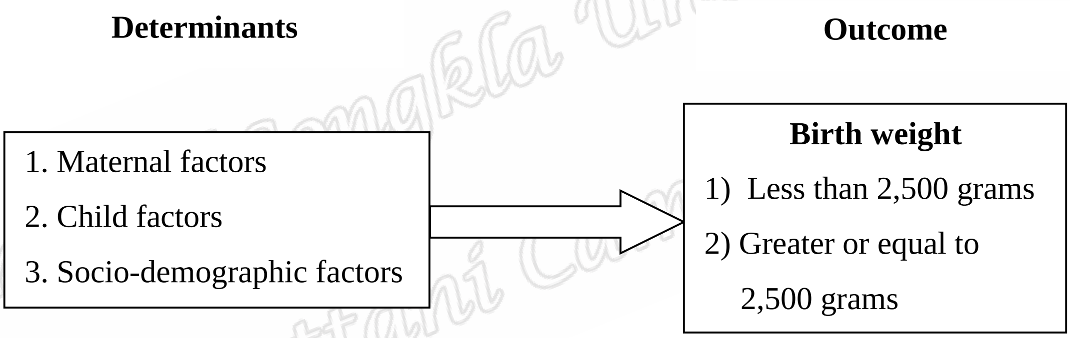

Figure 2.2 Path diagram of the study

2.5 Data management

2.5.1 Overall data management and its process

The overall data management and its process are presented in Figure 2.3.

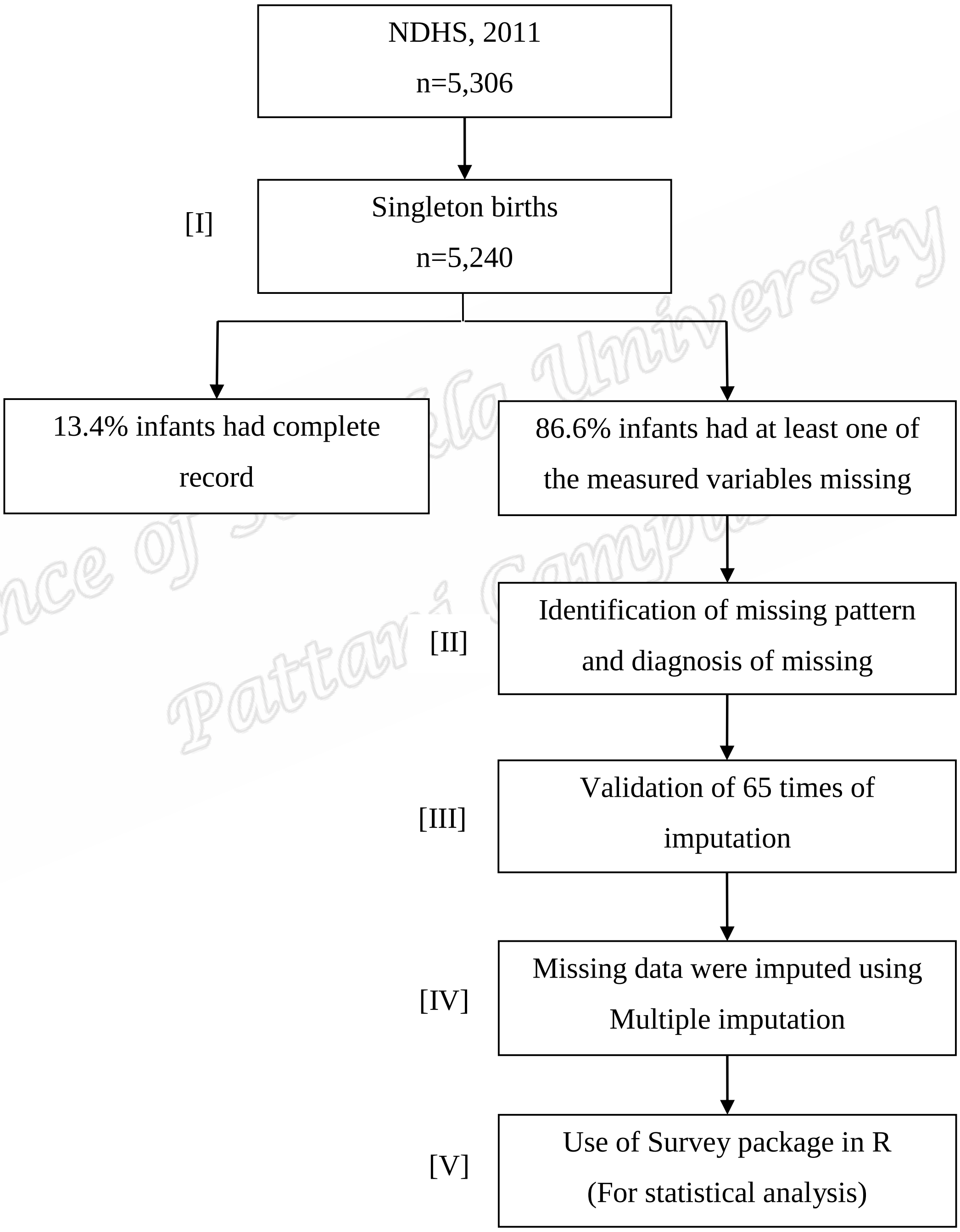

Figure 2.3 Flow diagram of overall data management

Figure 2.3 shows the overall data management and its process. The child data set contained 5,306 infants record. In the first step [I], 66 infants with multiple births were excluded, because they were likely to have LBW. However, out of 5,240 children, 13.4% children had complete record and 86.6% children had at least one of the measured variables missing. Hence, in the second step [II], missing data pattern and missing mechanism were diagnosed. After diagnosis of missing mechanism, number of times of imputation was validated in step [III]. Consequently, each missing value was imputed for 65 times using multiple imputation in step [IV]. Finally, survey package was applied to each imputed data set for statistical analysis.

2.5.2 Identification of frequency, pattern and missing mechanism of missing data

Out of 5,240 children, 696 (13.4%) children had complete record and 4,544 (86.6%) children had at least one of the measured variables missing. Therefore, to handle missing data, imputation was employed in this study. Nevertheless, before handling missing data, it was necessary to understand frequency, pattern and mechanism of missing data. The frequency and pattern of missing data was checked using package VIM in R.

To differentiate missing mechanism between MCAR and MAR statistically, Little's test was implemented (Little, 1988). To distinguish between MAR and MNAR mechanism there is no statistical test, which can be just reasoned or hypothesized (McKnight *et al.*, 2007).

2.5.3 Imputation

As aforementioned in the last section, missing data in this study held MAR assumption; therefore, this study aimed to impute missing birth weight rather than analyzing only the measured birth weight. There are several methods of imputation, but in this study multiple imputation was employed to account missing data, because multiple imputation provides better estimates of parameters than other methods, when missing data are MAR.

Multiple imputation

To overcome the limitations of single imputation, Rubin (1987) proposed multiple imputation method where each missing value is imputed by *m* > 1 times resulting into *m* data set. Each data set is analyzed by using complete data method.
The estimates of parameters of *m* data set are pooled to get overall estimates and standard errors that show uncertainty created by missing data.

For combining estimates of parameters of *m* dataset, formulas derived by Rubin (1987) is used. Suppose the regression coefficient for an imputed dataset *i* is Q_i and U_i be the variance where *i*= 1, 2, ..., m. Therefore, the overall regression coefficient is the average of all Q_i and shown in equation 1.

$$\bar{Q} = \frac{1}{m}\sum_{i=1}^{m} Q_i \quad (1)$$

The variance within imputation is average of all U_i and is shown in equation 2.

$$\bar{U} = \frac{1}{m}\sum_{i=1}^{m} U_i \quad (2)$$

The variance between imputaion is displayed in equation 3.

$$B = \frac{1}{m-1}\sum_{i=1}^{m} (Q_i - \bar{Q})^2 \quad (3)$$

The total variance is a combination of variance within and in between imputations which is displayed in equation 4.

$$T = \bar{U} + \left(1 + \frac{1}{m}\right) B \tag{4}$$

The overall standard error is the square root of total variance T and is displayed in equation 5.

$$SE = \sqrt{T} \tag{5}$$

In multiple imputation, methods like Joint Modeling (JM) and Multiple Imputation by Chained Equations (MICE) also called as Fully Conditional Specification (FCS) have been proposed to impute missing data (van Buuren and Groothuis-oudshoorn, 2011). In MICE approach, chains of regression models are performed in which each determinant with missing value is modeled conditionally upon other determinants in the data. This signifies that each variable has its own imputation model. For example, logistic regression model is used for dichotomous determinants and linear regression model used for continuous determinants (Azur *et al.*, 2011). As described by van Buuren and Groothuis-oudshoorn (2011), multiple imputation involves three main steps: imputation, analysis and pooling. Firstly, an imputation model is used to generate the missing values using possible values. In the imputation model, auxiliary variables and variables that can explain missing mechanism are kept for better prediction of missing values and making MAR hypothesis more possible (Schafer, 1997; Collins *et al.*, 2001). Rubin (1987) claimed that best result can be made when the number of imputation *m* ranges in between 3 to 5. However, Graham *et al.* (2007) suggested the use of number of imputations (*m*) of 20, 20, 40, 100 and above 100 for missing information (γ) of 0.10, 0.30, 0.50, 0.70 and 0.90 respectively in order to achieve the change in statistical power to be less than one percent. In spite of this, White *et al.* (2011) argued that along with statistical efficiency and power, Monte Carlo errors should be considered when setting the optimal number of imputations and suggested that number of imputations must be at least equal to percentage of missing data. Secondly, analysis model is applied to estimate parameters for each imputed data set. Basically, in theory, the analysis model and the imputation model need to be same, but they can be different in practice (Schafer, 1997). Finally, the estimated coefficients, standard errors and confidence intervals from each model are pooled together using Rubin's rule.

2.6 Implementation and Statistical analysis

The pattern of missing data in this study was arbitrary and missing variables were categorical; hence, FCS method was considered appropriate
(van Buuren and Groothuis-oudshoorn, 2011). Therefore, mice package was used in this study, because in mice package multiple imputation using FCS is implemented by MICE algorithm (van Buuren and Groothuis-oudshoorn, 2011). For combining each imputed data set, mitools package by Lumley (2006) was applied. In this study, multiple imputation was carried out for 65 times, because the highest percentage of missing was 63.3. For the validation of number of imputations, 63.3 percent of observations from the measured birth weight data set (1,922) were randomly removed and deemed missing. Subsequently, the missing values were imputed for 5, 10, 25, 50, 65, 75 and 100 times.

As suggested by Boerma *et al*. (1996) mother's opinion on infant's birth size can be used as an alternative to the birth weight. Therefore mother's opinion on infant's size at birth was considered an auxiliary variable in this study. Before the imputation model, possible interaction between the variables like ANC and iron tablets consumption during pregnancy, wealth index and mother's education, education and women's decision for utilization of health services, ecological region and developmental region, and development region and women's decision for utilization of health services was assessed. Transform-then-impute method as described by von Hippel (2009) was employed to check for interaction and found no interaction among them, because the *p*-value was greater than 0.05. Consequently, all the study variables along with auxiliary variable were included into the imputation model. Survey package by Lumley (2011) in R was applied to each imputed data sets to account for sampling method and sample weights. Survey logistic regression model as an analysis model was deployed to identify factors associated with LBW. For univariate analysis simple survey logistic regression was employed. For multivariate, all variables were included in multiple survey logistic regression model. Backward elimination was used to obtain the final model. Under complex survey data, the parameters are estimated by pseudo likelihood method instead of maximum likelihood (Lee and Forthofer, 2005). Therefore, the adjusted Wald test statistic was applied for selecting significant variables.

2.7 Logistic regression model

Survey logistic regression is also a generalized linear model (GLM) which gives a fitted regression model that is linear in the coefficients for the covariates (Heeringa *et al*., 2010). Suppose γ which is the outcome variable is binary with possible values 0 and 1. Where 0 represents babies with normal birth weight and 1 indicates babies with low birth weight. E $(\gamma|x) = \pi(x)$ is the conditional probability, denoting $\gamma = 1$ given the variable x. Therefore, $\pi(x)$ is probability that $\gamma = 1$, conditional on a vector of determinants x. $g(\pi(x))$ is a nonlinear function of $\pi(x)$ that results in a fitted regression model. The model is linear in the coefficients for the variables. The logit is the link function used to model dichotomous survey determinants and for a binary logistic regression model. The logit function is given by

$$g((x)) = \mathrm{logit}(\pi(x)) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \qquad (6)$$

where β_0 is constant, β_p are model coefficients and x_p are model variables.

The results are displayed in the form of odds ratio, which is obtained by exponentiation of coefficients.

$$OR = \exp(\beta) \qquad (7)$$

Chapter 3

Results

3.1 Frequency, pattern and missing mechanism of missing data

Figure 3.1 shows the frequency and pattern of missing data. The percentage of missing values in each variable and the pattern of missing data is displayed in Figure 3.1, left (a) and right (b) panel respectively. From the (a), the highest percentages of missing values were under the variables birth weight (63.3%), and BMI (52.0%). The percentages of missing were nearly equal for ANC and consumption of iron tablets during pregnancy, whereas the percentage of missing was less than 10 for cooking fuel and women's decision for utilization health of services. There were no missing values for the variables such as mother's age at child's birth, gender of child, parity, mother's education, wealth index, ethnicity, residence, ecological region, development region, birth interval and smoking. The pattern of missing data is shown in (b). From the figure (b), only 13.4 percent of children had complete record without any missing values, while 21.2 percent of children contained missing values only on birth weight, whereas 15.4 percent children contained missing values only on BMI. Furthermore, 21.9 percent of children had missing values on both birth weight and BMI, and only 8.7 percent of infants contained missing values on birth weight, mother's BMI, ANC visit and consumption of iron tablets during pregnancy.

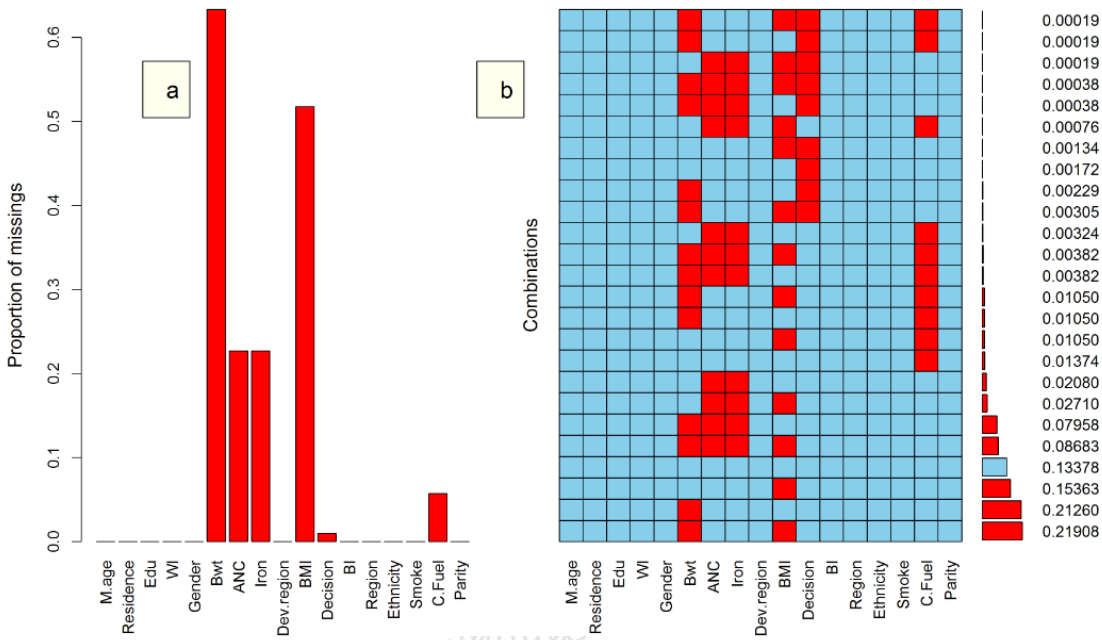

Figure 3.1 Percentage and pattern of missing data

Note: *M.age: Mother's age at child's birth, Edu: education, WI: wealth index, Bwt: birth weight, Iron: consumption of iron tablets during pregnancy, Decision: women's decision for utilization of health services, BI: birth interval and C.Fuel: cooking fuel.*

From the Little's test to distinguish between missing mechanisms, the *p*-value was found to be 0.000. Hence, missing mechanism was not MCAR. For diagnosis of missing mechanism between MAR and MNAR, reasoning was accounted. In this study, missing values on the birth weight were due to home delivery, because mothers who gave birth at home were possibly unable to provide numeric birth weight of their babies. Likewise, missing values on other dependent variables were not missing due to themselves, but were missing due to other characteristics. Hence, missing data held MAR assumption.

3.2 Validation of imputation

For the validation of imputation, complete measured birth weight data set (1,922) was used. This is because it was easy to compare estimates calculated after imputation with estimates obtained from complete data set. The standard error for the prevalence of LBW from the complete measured birth weight data set and the standard errors for the prevalence of LBW after imputing for seven different times were compared and shown in Figure 3.2. From the figure, it is seen that the standard error for prevalence of LBW increases with the increase in the number of imputations from 5 to 10 and then gradually decreases, when the number of imputation is increased to 25, 50 and 65 times. At 65 times, it can be seen that the standard error for the prevalence of LBW was lowest. However, after 65 times, standard error increases rapidly with the increase in number of imputations.

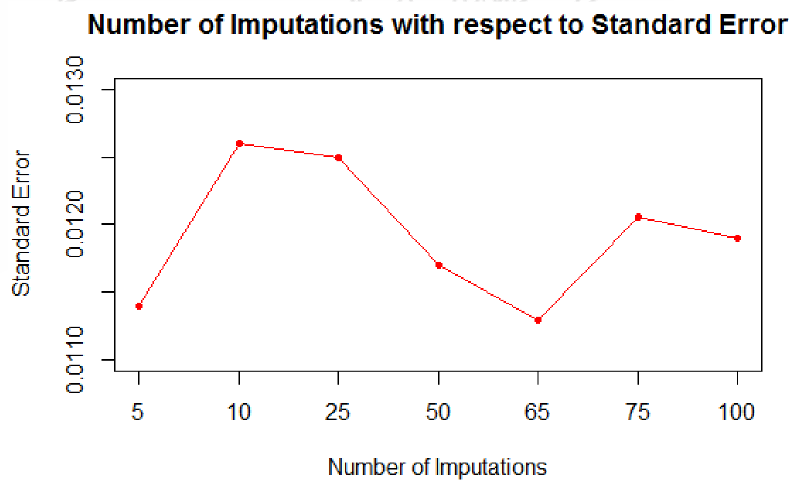

Figure 3.2 Number of imputations with respect to standard error

3.3 Data after imputation

The overall and crude subgroup estimations of LBW prevalence from the original data set and after multiple imputation were calculated and shown in

Table 3.1(a), 3.1(b) and 3.1(c). The overall prevalence of LBW only from the available birth weight was 11.5 percent and after imputation was 15.4 percent. The prevalence of LBW obtained from observed birth weight data set and after imputation were nearly equal in each subgroup for determinants such as mother's age at child's birth, gender of child, ethnicity and residence. However, the prevalences of LBW were different in each subgroup for the rest of the variables. Mothers, who attended primary education and whose decision for seeking health services relied only on husbands or others and gave birth to infants within a gap of less than 24 months from previous birth, showed the highest prevalence of having LBW babies compared to their respective subgroups in both observed and imputed data set. The percentage of giving birth to LBW infants obtained from measured birth weight data set was higher among mothers who had normal body weight (13.7%) and underweight (12.9%) than mothers who were overweight (8.0%). However, the percentage of giving birth to LBW infants obtained from measured birth weight data set was the highest among mothers who were underweight (17.6%) compared to respective subgroups. For ANC visit during pregnancy, the prevalence of giving birth to LBW infants calculated from measured birth weight data set was higher among mothers who attended one to three visits (11.4%) and four or more visits (10.5%) than mothers who did not attend ANC visit (8.0%). In contrast, the prevalence of giving birth to LBW infants obtained after imputation was highest among mothers who did not attend ANC visit (20.7%) compared to respective subgroups.

The percentages of LBW calculated from the measured birth weight data set for determinants such as iron tablet consumption during pregnancy, smoking, wealth index, cooking fuel and region were equal in each subgroup, whereas the percentages of LBW obtained after imputation for the above mentioned variables were different in each subgroup. In available birth weight data set, the higher prevalences for development region were evident in mothers residing in Eastern (15.0%), Far-western (14.2%) and Mid-western (12.3%) than mothers residing in other development regions. The prevalence of LBW for the development region obtained after imputation was the same, but the order was changed. The higher prevalences for development region were evident in mothers residing in Far-western (19.2%), Eastern (18.3%) and Mid-western (17.4%) than mothers residing in other development regions.

Table 3.1(a) Overall and subgroup prevalence of LBW

Variables	Measure birth weight	After imputation
Overall	11.5	15.4
Maternal factors		
Age at child's birth (Years)		
15-19	12.3	16.0
20-24	11.4	14.9
25-29	11.6	14.9
≥ 30	10.6	16.2
Education		
No education	9.0	16.1
Primary education	14.6	18.3
Secondary/ higher education	11.6	12.4
Body mass index (BMI)		
< 18.5 (Underweight)	12.9	17.6
18.5-23.0 (Normal)	13.7	16.0
> 23.0 (Overweight)	8.0	11.5
ANC visit during pregnancy		
No visit	8.0	20.7
One-three visits	11.4	16.7
Four or more visits	10.5	12.5
Consumption of iron tablets during pregnancy		
No	10.1	18.9
Yes	10.7	14.3
Smoking		
No	9.6	14.9
Yes	11.5	21.0

Table 3.1(b) Overall and subgroup prevalences of LBW

Variables	Measured birth weight	After imputation
Parity		
One	10.6	14.0
Two-three	13.0	15.4
Four and above	8.0	16.4
Women's decision for health service utilization		
Women	8.4	10.5
Women and husband together	11.8	15.2
Husband or others	13.4	18.1
Child factors		
Birth interval		
No interval	12.3	15.0
< 24 months	13.1	19.1
≥ 24 months	10.2	14.6
Gender of baby		
Male	10.7	14.6
Female	12.4	16.1
Socio-demographic factors		
Wealth index		
Poor	11.8	17.4
Middle	10.8	15.2
Rich	11.6	12.3
Ethnicity		
Relatively advantaged	12.2	15.9
Relatively disadvantaged (Janjati)	10.2	14.2
Relatively disadvantaged (Dalit)	10.9	16.0

Table 3.1(c) Overall and subgroup prevalences of LBW

Variables	Measured birth weight	After imputation
Fuel		
Low polluting fuel	10.3	11.0
High polluting fuel	11.4	16.1
Residence		
Rural	11.4	15.5
Urban	10.2	14.0
Ecological region		
Mountain	12.6	17.8
Hill	11.9	17.0
Terai	11.2	13.7
Development region		
Eastern	8.6	18.3
Central	15.0	12.1
Western	9.0	13.1
Mid-western	12.3	17.4
Far-western	14.2	19.2

3.4 Factors associated with LBW

For the univariate analysis, all study variables were analyzed by using simple survey logisitc regression and results are displayed in Table 3.2(a), 3.2(b) and 3.2(c). Women's decision for utilization of health services and cooking fuel was found statistically significant. Mothers had more tendency to give birth to LBW infants, when decision on utilization of health services relied only on husband and others (OR 1.91, 95% CI = 1.34, 2.72) and, mother and her husband together (OR 1.54, 95% CI = 1.03, 2.30). Mothers using high polluting cooking fuels (OR 1.56, 95% CI = 1.07, 2.28) were more inclined to give birth to LBW infants than mothers using non polluting cooking fuels. However, the variables like wealth index, mother's education, ethnicity, residence, ecological region, developmental region, mother's BMI, birth interval, ANC visit, consumption of iron tablets during pregnancy, smoking, mother's age at child's birth, parity and gender of child remain insignificant with LBW.

Table 3.2(a) Unadjusted odds ratio and 95% CI of study variables

Variables	Unadjusted OR	95% CI	*p*-value
Maternal factors			
Age at child's birth (Years)			0.970
≥ 30	1.00		
25-29	0.91	0.55, 1.51	
20-24	0.92	0.56, 1.50	
15-19	0.99	0.60, 1.66	
Education			0.107
Secondary/ higher education	1.00		
Primary education	1.58	1.09, 2.27	
No education	1.34	0.90, 2.00	
Body mass index (BMI)			0.138
> 23.0 (Overweight)	1.00		
18.5-23.0 (Normal)	1.50	0.96, 2.33	
< 18.5 (Underweight)	1.67	1.03, 2.71	
ANC visit during pregnancy			0.131
Four or more visits	1.00		
One-three visits	1.38	0.96, 1.98	
No visit	1.83	0.96, 3.51	
Consumption of iron tablets during pregnancy			0.199
Yes	1.00		
No	1.39	0.84, 2.30	
Smoke			0.247
No	1.00		
Yes	1.47	0.75, 2.88	

Table 3.2(b) Unadjusted odds ratio and 95% CI of study variables

Variables	Unadjusted OR	95% CI	*p*-value
Parity			0.748
Four and above	1.00		
Two-three	0.94	0.60, 1.48	
One	0.84	0.51, 1.38	
Women's decision for health service utilization			0.002*
Women	1.00		
Women and husband together	1.54	1.03, 2.30	
Husband or others	1.91	1.34, 2.72	
Child factors			
Birth interval			0.338
No interval	1.00		
< 24 months	1.32	0.82, 2.15	
≥ 24 months	0.96	0.69, 1.33	
Gender of baby			0.379
Male	1.00		
Female	1.13	0.86, 1.48	
Socio-demographic factors			
Wealth index			0.107
Rich	1.00		
Middle	1.27	0.83, 1.96	
Poor	1.50	1.03, 2.20	
Ethnicity			0.758
Relatively advantaged	1.00		
Relatively disadvantaged (Janjati)	0.88	0.61, 1.26	
Relatively disadvantaged (Dalit)	1.00	0.67, 1.49	

Table 3.2(c) Unadjusted odds ratio and 95% CI of study variables

Variables	Unadjusted OR	95% CI	*p*-value
Fuel			0.023*
Low polluting fuel	1.00		
High polluting fuel	1.56	1.07, 2.28	
Residence			0.466
Urban	1.00		
Rural	1.12	0.82, 1.53	
Ecological region			0.289
Terai	1.00		
Hill	1.29	0.89, 1.87	
Mountain	1.37	0.87, 2.17	
Development region			0.072
Central	1.00		
Eastern	1.63	1.03, 2.57	
Western	1.09	0.64, 1.84	
Mid-western	1.53	0.92, 2.57	
Far-western	1.72	1.07, 2.78	

p-value was calculated from Wald test, *statistically significant at 5% level.

For multiple logistic regression, all study variables were included in the model. The results from the full model are presented in Table 3.3(a), 3.3(b) and 3.3(c). The table shows the adjusted odds ration with its 95% CI and *p*-value from Wald test. The inference statistical tests obtained from the full model were changed compared to simple logistic regression model. From the result, it is seen that only women decision for the utilization of the health services remain significant determinant. Women with the lowest autonomy on their own health compared to those with involvement of husband or others (adjusted OR 1.91, 95% CI = 1.53, 2.28) and with husband and women together (adjusted OR 1.64, 95% CI = 1.23, 2.05) had more chance to give birth to LBW infants. However, cooking fuel which was significant in univariate analysis remained insignificant in the full model. Likewise, the variables like wealth index, mother's education, ethnicity, residence, ecological region, developmental region, mother's BMI, birth interval, ANC visit, consumption of iron tablets during pregnancy, smoking, mother's age at child's birth, parity and gender of child also remained insignificant with LBW.

A backward elimination procedure was employed for selecting best model. In this procedure, all the study variables were kept in the full model and then a variable with the highest p-value that did not meet significance (p-value > 0.05) was deleted from the model. This procedure was repeated until we got significant variables (p-value less than 0.05). From this procedure, only women's decision for the utilization of health services remained significant. Therefore, in the final model only women's decision for the utilization of health services was kept. In this model, mothers had more tendency to give birth to LBW infants, when decision on utilization of health services relied only on husband and others (OR 1.91, 95% CI = 1.34, 2.72) and mother and her husband together (OR 1.54, 95% CI = 1.03, 2.30).

Table 3.3 Adjusted odds ratio and 95% CI of study variables

Variables	Adjusted OR	95% CI	*p*-value
Maternal factors			
Age at child's birth (Years)			0.962
≥ 30	1.00		
25-29	0.98	0.46, 1.50	
20-24	0.90	0.36, 1.44	
15-19	0.89	0.30, 1.48	
Education			0.311
Secondary/ higher education	1.00		
Primary education	1.36	0.92, 1.80	
No education	1.02	0.54, 1.52	
Body mass index (BMI)			0.334
> 23.0 (Overweight)	1.00		
18.5-23.0 (Normal)	1.30	0.84, 1.76	
< 18.5 (Underweight)	1.55	1.03, 2.07	
ANC visit during pregnancy			0.320
Four or more visits	1.00		
One-three visits	1.32	0.94, 1.70	
No visit	1.76	0.90, 2.64	
Consumption of iron tablets during pregnancy			0.885
Yes	1.00		
No	0.95	0.32, 1.58	
Smoke			0.452
No	1.00		
Yes	1.29	0.63, 1.94	

Table 3.3(b) Adjusted odds ratio and 95% CI of study variables

Variables	Adjusted OR	95% CI	*p*-value
Parity			0.727
Four and above	1.00		
Two-three	1.16	0.66, 1.66	
One	1.02	0.33, 1.71	
Women's decision for health service utilization			0.006*
Women	1.00		
Women and husband together	1.64	1.23, 2.05	
Husband or others	1.91	1.53, 2.28	
Child factors			
Birth interval			0.317
No interval	1.00		
< 24 months	1.05	0.47, 1.64	
≥ 24 months	0.77	0.28, 1.27	
Gender of baby			0.385
Male	1.00		
Female	1.13	0.85, 1.41	
Socio-demographic factors			
Wealth index			0.977
Rich	1.00		
Middle	1.06	0.57, 1.55	
Poor	1.05	0.52, 1.57	
Ethnicity			0.492
Relatively advantaged	1.00		
Relatively disadvantaged (Janjati)	0.80	0.45, 1.17	
Relatively disadvantaged (Dalit)	0.87	0.42, 1.32	

Table 3.3(c) Adjusted odds ratio and 95% CI of study variables

Variables	Adjusted OR	95% CI	*p*-value
Fuel			0.613
Low polluting fuel	1.00		
High polluting fuel	1.15	0.62, 1.67	
Residence			0.352
Urban	1.00		
Rural	0.85	0.52, 1.18	
Ecological region			0.441
Terai	1.00		
Hill	1.29	0.93, 1.65	
Mountain	1.10	0.62, 1.58	
Development region			0.107
Central	1.00		
Eastern	1.75	1.27, 2.23	
Western	1.14	0.58, 1.71	
Mid-western	1.38	0.85, 1.90	
Far-western	1.69	1.18, 2.20	

p-value was calculated from Wald test, *statistically significant at 5% level.

Chapter 4

Discussion and Conclusions

4.1 Discussion

4.1.1 Imputation

In this study, missing values were presented in birth weight, BMI, ANC visit and consumption of iron tablets during pregnancy, mother's decision on utilization of health services and cooking fuel. Out of six variables, the highest percentages of missing values were under the variables birth weight (63.3%), and BMI (52.0%). The pattern of missing data shown in Figure 3.1 (b) was arbitrary, because the missing values for the variables of any record were seen in a random fashion. In this study missing data were not MNAR, because missing values in the variables were not missing due to themselves, but were missing due to other characteristics. For example, missing data on birth weight were due to home delivery, because mothers who gave birth at home were possibly unable to provide numeric birth weight of their babies. The reason for missing values on ANC visit was probably mothers living in rural areas felt uncomfortable to report their number of ANC visits during the time of interview. Missing values on consumption of iron tablets during pregnancy might be due to missing values on ANC visit, because mothers who did not report their ANC visit were less likely to report consumption of iron tablets during pregnancy. Missing values on BMI occurred because of refusal to measure height and weight either by respondent or respondent's mother. In the NDHS, 2011, women belonging to the household (de jure) and those who did not belong to the household (non de jure) were interviewed. The later were might be visitors who came to visit household's members. However, only those women who belonged to the household were included in the individual file such as child file, women file, etc. Therefore, missing values for cooking fuel were for non de jure mothers. The result from Little's test also showed that missing data are not MCAR. So, the only possibility of missing mechanism is MAR. Hence, in this study missing data held MAR assumption.

Multiple imputation is employed in this study to handle missing data, because analysis based on only complete cases of measured birth weight cannot be used since missing data are presented in more than one variable, and the missing data are MAR. Even though mother's recall of infant's size at birth weight is considered an alternative to the birth weight, but in this study imputation method provides a better option than using size at birth to birth weight, when missing data also occur in other variables. In this study, multiple imputation was carried for 65 times because the standard error for prevalence of LBW carried out from 65 times of imputation was lowest compared to other number of imputations as shown in Figure 3.2. Similarly, the standard error for the prevalence of LBW calculated from 65 times (0.0113) was closest to the standard error for the prevalence of LBW calculated from the complete observed infant's birth weight (0.0097). The findings of this study confirmed the findings of White *et al.* (2011) that used Monte Carlo error to identify the optimize number of imputations and found that the number of imputations should be at least equal to percentage of missing data.

The percentage of missing in this study is 63.3 and therefore, the statistical analysis is prone to bias (Bennet, 2001). However, the missing data pattern and missing mechanisms are more important in imputing missing than the percentage missing

(Tabachnick and Fidell, 2007). Nevertheless, using multiple imputation leads bias downwards compared to analysis of complete cases, but it does not mean that using imputation methods for replacing missing values removes the bias completely.

The current study consists of missing data on the variables like birth weight, BMI, ANC visit, consumption of iron tablets during pregnancy, cooking fuel and women's decision for utilization of health services. For birth weight, even though there has been a considerable rise in the percentage of measurement of infants birth weight at birth in the past five years from 17 in 2006 to 36 percent in 2011 (MOHP and New ERA, 2006; MOHP and New ERA, 2011), but women giving birth at home is ongoing problem as stated in (Karkee *et al*., 2014; Sreeramareddy *et al*., 2006). Unfortunately, the problem of missing data on birth weight will continue till future for low income countries. This suggests that promotion and strengthening institutional delivery and provision of weighing scale and training to community health workers should be implemented for measurement of birth weight of those infants who are born at home. However, missing data available in other variables can be minimized. For instance, in DHS survey, mothers belonging to the household (de jure) and those who did not belong to the household (non de jure) were interviewed. However, questions related to cooking fuel were collected in household level and assigned to individuals in the individual data file. Thus, a mother who was not member of household lacked the data on cooking fuel. The problem of missing data on cooking fuel can be avoided by including questions related to cooking fuel in women's questionnaire or by not interviewing a mother who is not the member of household.

4.1.2 Factors associated with LBW

The overall prevalence of LBW from this study was 15.4 percent which is different from the study including only infants with measured birth weight conducted by Khanal *et al*. (2014) in which the prevalence of LBW was found 11.5 percent. The difference is expected, because in this study there is an inclusion of additional 3,318 missing birth weight in the analysis. It is also suggested that the prevalence of LBW obtained from analyzing only infants with measure birth weight is underestimated. In a study conducted by Khanal *et al*. (2014), it is found that prevalence of small size at birth is 16 percent, and close to the prevalence of this study. This may be because mother's recall of infant's size at birth and other variables are used for imputing missing values in this study. As shown in Table 3.1(a), 3.1(b) and 3.1(c), the prevalence of LBW obtained from observed birth weight data set and after imputation is nearly equal in each subgroup for the determinants such as mother's age at child's birth, gender of child, ethnicity and residence. It is suggested that each subgroup has equal chance of having LBW infants. In this study, the prevalence of having LBW infants was higher among mothers living in low standard such as poor, used high polluting cooking fuels, did not attend ANC visit and did not consume iron tablets during pregnancy compared to their respective subgroups and this finding was consistent with previous study conducted by (Khanal *et al*., 2014). For no ANC visit and mothers who smoke, the confidence intervals are wider compared other subgroups. The reason is possibly due to small sample size in these two categories.

The current study found that mothers had more affinity to give birth to LBW babies, when decision on utilization of health services is relied only on others instead of themselves and this finding is supported by Sharma and Kader (2013) in which women with lowest decision making autonomy were more likely to have LBW. This is probably because women with lowest decision making autonomy on their health care are less tended to get regular health checkups together with ANC visit during pregnancy including safe deliveries and health information regarding pregnancy and childbirth. Apart from that, women with lowest decision making autonomy on their own health may have poor nutrition uptake during pregnancy and which may consequently impair fetal growth. Variables such as ANC visit during pregnancy and consumption of iron tablets during pregnancy are not significant with LBW in the current study. However, studies conducted in Nepal found that mothers who did not attend ANC visit during pregnancy and mothers who did not consume iron tablets during pregnancy were more inclined to give birth to LBW infants (Khanal *et al*., 2014; Khanal *et al*., 2014). This may be because, the above mentioned studies assumed missing values present on ANC visit and iron tablets consumption during pregnancy as no ANC visit and no consumption of iron tablets during pregnancy, while in this study both missing values presented in ANC visit and consumption of iron tablets during pregnancy were imputed.

4.2 Conclusion

In this study, multiple imputation was implemented to overcome missingnes presented on both determinants and outcome, ranging from 1 percent to 63 percent. The result of this study showed that one in every six infants had LBW after imputation. The mothers with the lowest autonomy on making decisions for utilization of health services had more affinity to give birth to LBW infants. This suggests that there is a need for implementing programs that focus on promotion and strengthening women autonomy as well as educational interventions not only to the couple but also to their in-laws regarding the importance of utilization of health services to help reduce the risk of birth to LBW infants among mothers

4.3 Limitation and suggestions for further study

Although, the missing data were MAR, the level of missingness was still very high (63.32%) in this study. The results obtained from this study were not consistent with the previous studies conducted in Nepal using the same survey data. Hence, to confirm results, it is suggested that the analysis should be repeated, when the level of missing decreases considerably. Furthermore, the current study utilized the secondary data; thus, exact reasons for missing data were not clear for many variables and they could have effect on the results.

There are more multiple imputation methods and other imputation techniques for handling missing data. For example, joint modeling, regression imputation, mean substitution and hot deck. Therefore, further study is needed to compare the methods of imputations for their efficiency in handling missing data.

The findings of this study suggested that obtaining the prevalence of LBW from only the sample of measured birth weight results in under estimation. In addition, assuming missing values as non missing provided different results from imputed. Therefore, it is suggested for future researchers conducting studies on LBW with DHS data from developing countries that missing data on birth weight and its determinants should be handled.

References

Almond, D., Chay, K.Y. and Lee, D.S. 2005. The costs of low birth weight. The Quaterly Journal of Economics, 120(3), 1031-1084.

Almond, D., Chay, K.Y., Lee, D.S. and Berkeley, U. C. 2002. "Does Low Birth Weight Matter ? Evidence from the U.S. Population of Twin Births", (Working Paper No. 53). Retrieved from Centre for Labor Economics website: http://cle.berkeley.edu/working_papers.shtml#4.

Azur, M.J., Stuart, E.A., Frangakis, C. and Leaf, P.J. 2011. Multiple imputation by chained equations : what is it and how does it work ? Intenational Journal of Mehods in Psychiatric Research, 20(1), 40–49.

Badshah, S., Mason, L., McKelvie, K., Payne, R. and Lisboa, P.J. 2008. Risk factors for low birthweight in the public-hospitals at Peshawar, NWFP-Pakistan. BMC Public Health, 8 (1), 197.

Balci, M.M., Acikel, S. and Akdemir, R. 2010. Low birth weight and increased cardiovascular risk: fetal programming. International Journal of Cardiology, 144(1), 110–1.

Bennett, D.A. 2001. How can I deal with missing data in my study? Australian and New Zealand Journal of Public Health, 25(5), 464–469.

Bethlehem, J. 2009. Applied survey methods: A statistical perspective. John Wiley and Sons, Hoboken, New Jersey. Doi: 10.1002/9780470494998.ch13.

Blanc, A.K. and Wardlaw, T. 2005. Monitoring low birth weight : an evaluation of international estimates and an updated estimation procedure. Bulletin of the

World Health Organization, 83(3), 178–185.

Boerma, J.T., Weinstein, K.I., Rutstein, S.O. and Sommerfelt, A.E. 1996. Data on birth weight in developing countries: can surveys help? Bulletin of the World Health Organization, 74(2), 209.

Bondevik, G.T., Lie, R.T., Ulstein, M. and Kvale, G. 2001. Maternal hematological status and risk of low birth weight and preterm delivery in Nepal. Acta Obstetricia et Gynecologica Scandinavica, 80 (5), 402–408.

Channon, A.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(05), 555–573.

Chiolero, A., Bovet, P. and Paccaud, F. 2005. Association between maternal smoking and low birth weight in Switzerland: the EDEN study. Swiss Medical Weekly, 135(35-36), 525–530.

Collins, L. M., Schafer, J.L. and Kam, C.M. 2001. A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychological Methods, 6(4), 330.

De Waal, T., Pannekoek, J. and Scholtus, S. 2011. Handbook of statistical data editing and imputation. John Wiley & Sons, Hoboken, New Jersey. Doi: 10.1002/9780470904848.ch7.

Dharmalingam, A., Navaneetham, K. and Krishnakumar, C.S. 2010. Nutritional status of mothers and low birth weight in India. Maternal and Child Health Journal, 14(2), 290–8.

Doherty, D.A., Magann, E.F., Francis, J., Morrison, J.C. and Newnham, J.P. 2006. Pre-pregnancy body mass index and pregnancy outcomes. International Journal of Gynaecology and Obstetrics, 95(3), 242–7.

Dong, Y. and Peng, C.Y.J. 2013. Principled missing data methods for researchers. SpringerPlus, 2(1), 1–17.

Fan, Z.J., Lackland, D.T., Lipsitz, S.R. and Nicholas, J. 2006. The association of low birthweight and chronic renal failure among medicaid young adults with diabetes and / or hypertension. Public Health Reports, 121(3), 239-244.

Gibbs, C.M., Wendt, A., Peters, S. and Hogue, C.J. 2012. The impact of early age at first childbirth on maternal and infant health. Paediatric and Perinatal Epidemiology, 26(s1), 259-284.

Graham, J.W., Olchowski, A.E. and Gilreath, T.D. 2007. How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. Prevention Science, 8(3), 206-213.

Heering, S.G., West, B.T. and Berglund, P.A. 2010. Applied Survey Data Analysis. CRC Press. USA., pp. 229-322.

Kalk, P., Guthmann, F., Krause, K., Relle, K., Godes, M., Gossing, G., Halle, H., Wauer, R. and Hocher, B. 2009. Impact of maternal body mass index on neonatal outcome. European Journal of Medical Research, 14(5), 216–222.

Karkee, R., Lee, A.H. and Khanal, V. 2014. Need factors for utilisation of institutional delivery services in Nepal: an analysis from Nepal Demographic and Health Survey, 2011. BMJ Open, 4(3), e004372.

Kayode, G.A., Amoakoh-Coleman, M., Agyepong, I.A., Ansah, E., Grobbee, D.E. and Klipstein-Grobusch, K. 2014. Contextual risk factors for low birth weight: a multilevel analysis. Public Library of Science One, 9(10), e109333.

Kent, S.T., McClure, L.A., Zaitchik, B.F. and Gohlke, J.M. 2013. Area-level risk factors for adverse birth outcomes: trends in urban and rural settings. BMC Pregnancy and Childbirth, 13(1), 129.

Khanal, V., Sauer, K., Karkee, R., and Zhao, Y. 2014. Factors associated with small size at birth in Nepal: further analysis of Nepal Demographic and Health Survey 2011. BMC Pregnancy and Childbirth, 14(1), 32.

Khanal, V., Zhao, Y. and Sauer, K. 2014. Role of antenatal care and iron supplementation during pregnancy in preventing low birth weight in Nepal: comparison of national surveys 2006 and 2011. Archives of Public Health, 72(1), 4.

King, J.C. 2003. The risk of maternal nutritional depletion and poor outcomes increases in early or closely spaced pregnancies. The Journal of Nutrition, 133(5), 1732S–1736S.

Kramer, M.S. 1987. Determinants of low birth weight: methodological assessment and meta-analysis. Bulletin of the World Health Organization, 65(5), 663.

Lawn, J.E., Cousens, S., and Zupan, J. 2005. 4 million neonatal deaths: When? Where? Why? The Lancet, 365(9462), 891-900.

Lee, E.S. and Forthofer, R.N. 2005. Analyzing complex survey data. Sage

Publications, Thousand Oaks, California. Doi: 10.4135/9781412983341.

Lira, P.I., Ashworth, A. and Morris, S.S. 1996. Low birth weight and morbidity from diarrhea and respiratory infection in northeast Brazil. The Journal of Pediatrics, 128(4), 497–504.

Little, R.J.A. 1988. A test of missing completely at random for multivariate data with missing values. Journal of the American Statistical Association, 83(404), 1198-1202.

Little, R.J.A. and Rubin, D.B. 2002. Statistical Analysis with Missing Data. Wiley Series in Probability and Statisitc, New York, USA. Doi:10.1002/9781119013563.

Lumley, T. 2006. mitools: Tools for multiple imputation of missing data. Retrieved from: URL: http://CRAN. R-Project. Org.

Lumley, T. 2011. Complex surveys: a guide to analysis using R. John Wiley & Sons, Hoboken, New Jersey.

McKnight, P.E., McKnight, K.M., Sidani, S. and Figueredo, A.J. 2007. Missing data: A gentle introduction. Guilford Press, New York. Doi: 10.1177/1098214007306655.

MOHP and New ERA. 2006. Nepal Demographic and Health Survey (NDHS) 2006. Kathmandu, Nepal: Ministry of Health and Population, New Era, and ICF International, Calverton, Maryland.

MOHP. 2011. Annual Report: Departmen of Health Services 2066/67 (2009/2010). Kathmandu, Nepal: Ministry of Health and Population.

MOHP and New ERA. 2011. Nepal Demographic and Health Survey (NDHS) 2011. Kathmandu, Nepal: Ministry of Health and Population, New Era, and ICF International, Calverton, Maryland.

Mondal, B. 1998. Low birth weight in relation to sex of baby, maternal age and parity: a hospital based study on Tangsa tribe from Arunachal Pradesh. Journal of the Indian Medical Association,96(12), 362-364.

Mondal, B. 2000. Risk factors for low birth weight in Nepali infants. Indian Journal of Pediatrics, 67(7), 477-482.

Muula, A.S., Siziya, S. and Rudatsikira, E. 2011. Parity and maternal education are associated with low birth weight in Malawi. African Health Sciences, 11(1), 65–71.

Nisar, Y.B. and Dibley, M.J. 2014. Antenatal iron-folic acid supplementation reduces risk of low birthweight in Pakistan: secondary analysis of Demographic and Health Survey 2006-2007. Maternal & Child Nutrition, 1–14.

Pedersen, M.G., Mortensen, P.B. and Webb, R.T. 2013. Birth weight, schizophrenia, and adult mental disorder: is risk confined to the smallest babies? Archives of General Psychiatry, 67(9), 923-930.

Reichman, N.E. and Pagnini, D.L. 1997. Maternal age and birth outcomes: data from New Jersey. Family Planning Perspective, 29(6), 268–272.

Robles, A. and Goldman, N. 1999. Can accurate data on birthweight be obtained from health interview surveys? International Journal of Epidemiology, 28(5), 925–931.

Rubin, D.B. 1976. Inference and missing data. Biometrika, 63(3), 581–592.

Rubin, D.B. 1987. Multiple imputation for nonresponse in surveys. John Wiley & Sons, Hoboken, New Jersey. Doi: 10.1002/9780470316696.ch1.

Schafer J.L. 1997. Analysis of incomplete multivariate data. Chapman and Hall, New York, USA. Doi: 10.1201/9781439821862.

Schafer, J.L. and Graham, J.W. 2002. Missing data: Our view of the state of the art. Psychological Methods,7(2), 147–177.

Sharma, A. and Kader, M. 2013. Effect of Women's Decision-Making Autonomy on Infant's Birth Weight in Rural Bangladesh. ISRN Pediatrics, 2013:159542

Silva, A.A.M., Bettiol, H., Barbieri, M.A., Pereira, M.M., Brito, L.G.O., Ribeiro, V.S. and Aragão, V.M.F. 2005. Why are the low birthweight rates in Brazil higher in richer than in poorer municipalities? Exploring the epidemiological paradox of low birthweight. Paediatric and Perinatal Epidemiology, 19(1), 43–49.

Silvestrin, S., Silva, C., Hirakata, V.N., Goldani, A.S., Silveira, P.P. and Goldanie, M.Z. 2013. Maternal education level and low birth weight: a meta-analysis. Jornal de Pediatria (Versão Em Português), 89(4), 339–345.

Sreeramareddy, C.T., Joshi, H.S., Sreekumaran, B.V., Giri, S. and Chuni, N. 2006. Home delivery and newborn care practices among urban women in western Nepal: a questionnaire survey. BMC Pregnancy and Childbirth, 6(1), 27.

Sreeramareddy, C.T., Shidhaye, R.R. and Sathiakumar, N. 2011. Association between biomass fuel use and maternal report of child size at birth-an analysis of

2005-06 India Demographic Health Survey data. BMC Public Health, 11(1), 403.

Sterne, J.A., White, I.R., Carlin, J.B., Spratt, M., Royston, P., Kenward, M.G. and Carpenter, J.R. 2009. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ, 338, b2393.

Sullivan, T.R., Salter, A.B., Ryan, P. and Lee, K.J. 2015. Bias and Precision of the "Multiple Imputation, Then Deletion" Method for Dealing With Missing Outcome Data. American Journal of Epidemiology, 182(6), 528–534.

Tabachnick, B.G. and Fidell, L.S. 2007. Using Multivariate Statistics. Allyn and Bacon, Boston, US. Doi:10.1037/022267.

UNICEF. 2014. Monitoring the situation of children and women. Retrieved May 17, 2015, from http://data.unicef.org/nutrition/low-birthweight

UNICEF and WHO. 2004. Low Birthweight: Country, regional and global estimates.New York: UNICEF and WHO.

Upadhyay, P., Liabsuetrakul, T., Shrestha, A.B. and Pradhan, N. 2014. Influence of family members on utilization of maternal health care services among teen and adult pregnant women in Kathmandu, Nepal: a cross sectional study. Reproductive Health, 11(1), 92.

van Burren, S. and Groothuis-oudshoorn, K. 2011. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1-67.

van Buuren, S. 2012. Flexible imputation of missing data. CRC press, Boca Raton, FL., pp. 7-9.

van der Heijden, G.J., Donders, A.R., Stijnen, T. and Moons, K.G. 2006. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. Journal of Clinical Epidemiology, 59(10), 1102–9.

von Hippel, P.T. 2007. Regression with missing Ys: An improved strategy for analyzing multiply imputed data. Sociological Methodology, 37(1), 83–117.

von Hippel, P.T. 2009. How To Impute Squares, Interactions, and Other Transformed Variables. Sociological Methodology, 39, 265–291.

Wang, X., Ding, H., Ryan, L. and Xu, X. 1997. Association between air pollution and low birth weight: A community-based study. Environmental Health Perspectives, 105(5), 514.

White, I.R., Royston, P. and Wood, A.M. 2011. Multiple imputation using chained equations: Issues and guidance for practice. Statistics in Medicine, 30(4), 377–99.

Wilcox, A. J. 2001. On the importance-and the unimportance-of birthweight. International Journal of Epidemiology, 30(6), 1233–1241.

WHO. 2004. International statistical classification of diseases and related health problems (Tenth., Vol. 1). Geneva, Switzerland: World Health Organization.

Appendix

Certificate of Appreciation

This is to certify that

Usha Singh

has presented an academic paper at

"The 4th International Conference"Global Health, Society and Human Development"

28th – 29th September 2015

Mahidol University, Salaya Campus, Thailand

Assistant Professor Dr. Penchan Pradubmook-Sherer
Head of Department of Society and Health
Faculty of Social Sciences and Humanities
Mahidol University

Estimating prevalence and factors associated with low birth weight in Nepal: further analysis of Nepal Demographic and Health Survey 2011

Usha Singh*, Attachai Ueranantasun, and Metta Kuning*****

attachai.u@psu.ac.th

***Master Student, Prince of Songkla University, Pattani, Thailand**

****Lecturer, Prince of Songkla University, Pattani, Thailand**

*****Associate Professor, Prince of Songkla University, Pattani, Thailand**

Abstract

To estimate prevalence of low birth weight (LBW) in developing countries is challenging due to highly missing measured birth weights. A study conducted by using only measured birth weight may depict good health status of children and women, and results may be biased. Therefore, this study is aimed to estimate missing birth weights and to identify the factors associated with LBW.

This study uses child dataset that contains information on under-five year children from Nepal Demographic and Health Survey (NDHS), conducted in 2011. Because most of mothers in the survey are having home delivery, measured weights for their children are missing due to unavailability of a measuring scale. Therefore, approximately 63% of overall birth weights are missing from the survey. To replicate these missing data, a linear regression model is used for imputation to obtain missing birth weights. Variables like mother's recall on size at birth, mother's age at child's birth, sub-region, and child's gender are applied for achieving imputation results. The available and replicated birth weights are consequently employed as a dependent variable. For identifying the association between LBW and independent variables, a univariate and multivariable logistic regression model is implemented by using R software.

The results reveal that a total of 5,240 children are analyzed in this study after excluding 66 multiple births. Infants with measured birth weight are 1,922 (37%), including 235 children (12%) with LBW (< 2,500grams), and this contribute to the mean birth weight of 3,030 grams. Of these estimated 3,318 (63%) birth weights; 556 (16%) have LBW

and mean birth weight is 2,983 grams. The estimated birth weights from the imputation are then combined with the measured birth weights. The combined result shows that 791 infants (15%) have LBW and consequently, the mean birth weight is 3,000 grams. The mother's who give birth during their teenage are more likely to give birth to LBW babies (odds ratio (OR) 1.47). Mother's who are uneducated are likely to deliver LBW babies (OR 1.24). Female infants are more likely to have LBW than male infants. Mother's residing in Western mountain; Far-western hill and Far-western terrain are likely to give LBW infants (OR 1.39, 1.67 and 1.42).

Introduction

Birth weight determines an infant's survival which means the lesser the weight of an infant, the greater the risk of an infant's dying (Wilcox, 2001). Infants whose weight at birth is less are twenty times riskier to die than normal weight infants (UNICEF & WHO, 2004). Therefore, birth weight is classified as low birth weight (LBW), when it is lower than 2,500 grams, and normal birth weight, when it is equal to or higher than 2,500 grams (World Health Organization, 2004). Similarly, LBW infants are more likely to suffer from diseases such as type II diabetes mellitus, heart diseases (Balci et al., 2010), psychological disorder (Pedersen et al., 2013) and kidney failure (Fan et al., 2006) during adult stage. As a result, there is significant economic burden to the society from LBW (Almond et al., 2005).

Globally more than 20 million babies are born every year, out of which 15.5 percent of all births represent as LBW and 95.6 percent of all LBW infants are born in developing countries (UNICEF & WHO, 2004). Several studies are conducted on LBW by using demographic and health survey (DHS) data and have found that factors like gender of the child (Khanal et al., 2014; Sreeramareddy et al., 2011), birth order, hemoglobin level, cooking fuel (Sreeramareddy et al., 2011), mother's education (Muula et al., 2011), antenatal visit during pregnancy, iron and folic acid supplementation during pregnancy (Khanal et al., 2014; Nisar & Dibley, 2014), geographical region (Khanal et al., 2014), residence and wealth index (Kayode et al., 2014) are associated with LBW.

However, Demographic Health Survey (DHS) data on birth weight are incomplete in the developing countries because birth weights of most infants are not measured due to home delivery (Blanc & Wardlaw, 2005). For this reason, it is suggested that mother's recall on birth size can be used as an alternative to birth weight (Blanc & Wardlaw, 2005; Boerma et al., 1996). Birth weight is considered the main indicator of neonatal and infant health (Almond et al., 2002). Studies that are conducted in developing countries have identified determinants that contribute small size at birth not on LBW (Khanal et al., 2014; Sreeramareddy et al., 2011). Nepal is one of the developing countries (World Bank, 2015), and correct estimation of percentage of LBW and its determinants is necessary for intervening programs in reduction of infant and neonatal mortality. NDHS (2011) has reported that, in Nepal, more than 50 percent of child's birth weight is not measured at the time of delivery because of home delivery. From the same survey report, the prevalence of LBW in Nepal is 11.5 percent (NDHS, 2011), which is not the exact prevalence of LBW. Estimating prevalence of LBW and identifying factors associated with it from the measured birth weight may result in to bias (Robles & Goldman, 1999). Thus, this study aims to estimate missing birth weights and to identify factors that are associated with low birth weight by using data from Nepal Demographic and Health Survey (NDHS) in 2011.

Literature review

Data on LBW

In many developing countries, it is difficult to identify prevalence of LBW because most of the deliveries take place at home and birth weights are not measured (Blanc & Wardlaw, 2005). Therefore, a study conducted by using hospital records to estimate LBW may not represent nationally. As an alternative, DHS survey began to collect information on birth weight based on both birth certificate or mother's recall (Blanc & Wardlaw, 2005). However, this type of survey data became limited because most of the infants weights are not measured at the time of birth. In 1996, the study showed utilization of mother's recall on size at birth (i.e. very large, larger than average, average, smaller than average and very small) as an alternative to the birth weight for estimating low birth weight (Boerma et al., 1996).

The study conducted to explore accuracy of birth weight information from six DHS survey data has revealed that estimating low birth only from the measured birth weight sample from the surveys of developing countries may portray an overly optimistic picture of children's and mother's health status (Robles & Goldman, 1999). The findings of this study is supported by a similar study which analyzed 62 DHS surveys of 42 developing countries; found mothers who delivered their babies at hospital were educated and lived in urban areas(Blanc & Wardlaw, 2005).

Determinants

The study conducted in US, shows U shaped relation between the mother's age and the LBW among whites. The results from this study suggest that mothers with the age below 15 and above 40 are more likely to deliver LBW babies than mothers between age 25-29 years (Reichman & Pagnini, 1997). However, a systematic review reveals an evidence of a dose-response relationship between mother's age at birth and LBW which means the degree of association is decreased with an increase in mother's age (Gibbs et al., 2012).

Several researchers have found an inverse relationship between mother's education and LBW. The study on LBW reports that the odds of delivering LBW is higher in those women who have no formal education compared to those who have at least secondary level of education (Muula et al., 2011). Another study in Nepal also shows that babies with LBW are significantly higher among illiterate mothers than literate ones (Mondal.B, 2000).

A multilevel analysis on contextual risk factors for low birth weight reports that chance of LBW is about 43% among mother living in rural areas which may be due to inadequate health facilities, lack of job opportunities and infrastructure development (Kayode et al., 2014). In contrast to this finding, the study from Brazil reports that prevalence of LBW is higher in city dwellers than rural dwellers, because of incomplete registration of live births and registration of live births as stillbirths in rural areas (Silva et al., 2005).

Several studies have proven that economic disadvantage is associated with low birth weight. For example, a study conducted in Ghana illustrates that women living in low economic status have the risk of having LBW infants twice than women living in wealthier communities (Kayode et al., 2014). The finding is consistent with a previous study conducted in Malawi which found that risk of delivering LBW infant is lower for mother having high economic status (Muula et al., 2011).

Several studies have revealed that female infants are born with LBW than male infants. The study conducted in India shows that incidence of LBW is greater among female babies compared to male babies (Mondal, 1998). Likewise, the study conducted in Nepal also reports that risk of being small in size is less among male infant than female infant (Khanal et al., 2014).

Methodology

Child dataset that includes informations of under five year children from Nepal Demographic and Health Survey (NDHS), conducted in 2011 is used in this study. NDHS survey is a nationally representative, in which multistage sampling is used (NDHS, 2011). In first stage, probability proportionate to size is used to select enumeration areas (EA's). An EA is a ward in rural areas and subward in urban areas. In second stage, random sampling is used to select households.

Definition of variables

Outcome variable

Birth weight of an infant is considered an outcome of this study and is divided into normal birth weight, equal to or greater than 2,500 grams, and LBW, lower than 2,500 grams.

Exposure variables

Variables consisting of mother's age at child's birth, mother's education, gender of a child and sub-region are employed as determinants in this study. Besides birth weight, NDHS 2011 asked a specific question to the mothers about the size of their babies at the time of birth. Based upon five categories namely very large; large; normal; small and very small, mothers had to recall their babies' size. This variable is also used as an auxiliary variable while imputing birth weight. Mother's age at child's birth is changed from continuous to categorical variable. Mother's age at child's birth is classified as 15-19 years, 20-24 years, 25-29 years and 30 years of age and over. Mother's education is categorized into no education, primary education, secondary education and higher education. Initially wealth index has five categories which are regrouped into three categories, including poor, middle and rich. Child's gender is classified as male and female. Sub-region is divided into thirteen categories, comprising Eastern Mountain, Central Mountain, Western Mountain, Eastern Hill, Central Hill, Western Hill, Mid-Western Hill, Far-western Hill, Eastern Terrain, Central Terrain, Western Terrain, Mid-western Terrain and Far-western Terrain.

Imputation

The child dataset contains information of 5,306 children who were born during period from 2006-2011. 33 pairs of children are twin babies. They are excluded from the study because twin babies are usually born with low weight (Khanal et al., 2014). Therefore, only 5,240 children are analyzed in this study. Out of 5,240 children, only 1,922 children have their birth weight reported in the dataset. Since 3,318 children birth weight are missing, an estimation of prevalence of LBW would be biased by using only reported birth weight (Robles & Goldman, 1999). Thus, this study attempts to replace missing birth weight by means of regression imputation.

Imputation is a method to predict and substitute an appropriate value for a missing value in a dataset (De Waal et al., 2011). In a regression imputation method, regression equation is used to estimate missing value y_i in *y* variable in data *i*, from *x* variable which contains complete data where the number of *x* variable ranges from zero to many (De Waal et al., 2011). Hence, the linear regression equation will be as follows (De Waal et al., 2011):

$$\tilde{y}_i=\hat{y}_i=\hat{\alpha}+x_i^T\hat{\beta}+\mathcal{E}_i \qquad \text{with residual}$$

$$\tilde{y}_i=\hat{y}_i=\hat{\alpha}+x_i^T\hat{\beta} \qquad \text{without residual}$$

where α and β are parameters of estimate, $\tilde{y}_i$ is an imputed value and Ɛ is a residual.

Bethlehem (2009) has stressed that an auxiliary variable that has strong relationship with an outcome variable should be used for imputation. Therefore, mothers recall on babies' size is used as an auxiliary variable because it is statistically found from a preliminary logistic regression that birth weight and mother's opinion on baby's size has strong relationship as *p*-value is less than 0.05. Variables like gender of the baby, mother's age at child's birth and sub-region also have relationship with birth weight ($p < 0.05$). Hence, they are also included in the regression imputation model for good result.

Before imputation, the consistency of birth weight in each category of mother's opinion on birth size from a subsample of measured birth weight is checked. It is found that, 213 (11%) of mothers recalled incorrectly about their babies' size with respect to birth weight. Initially, it is assumed that mothers those who are uneducated, live in rural areas and poor may report incorrectly birth size. In order to see if the significant difference in socio-economic status has an effect on mother’s opinion and thus prove the assumption, mother's opinion on birth size is classified into 0 (correctly reported i.e.

1,709) and 1 (incorrectly reported i.e. 213). Thereafter, logistic regression model is fitted with wealth index, mother's education and residence as independent variables and mother’s opinion on child's size as a dependent variable. The result from logistic regression shows that none of the variables is significant. Hence, mother's opinion on birth size is able to be used as an auxiliary variable for imputation because it is not biased by socio-economic status. After imputation, both measured birth weight and imputed birth weight are subsequently combined together to be used for analysis.

Statistical analysis

Prevalence of LBW is calculated from descriptive statistics. Chi-square test is used to identify association between birth weight and independent variables. Multiple logistic regression is used to identify the strength of association between birth weight and independent variables. Variables that are significant in the Chi-square test are only included into the multiple logistic regression. Backward elimination is used for selection of variables in final model. For the model assessment, Pearson goodness of fit statistics is used. All statistical analyses are carried out by using R software version 3.1.3.

Ethics

Ethical approval for the conduction of NDHS survey is obtained by Nepal Health Research Council, Nepal and ICF Macro International Review Board in Calverton, Maryland, USA; and the data analysis protocol is approved by the Curtin University Human Research Ethics Committee.

Results

Birth weight

A total of 5,240 children are analyzed in this study. The description of an outcome is shown in Table 1. Out of 5,240 infants, 3,318 infants birth weight are imputed because of missing birth weight. From the imputed birth weight, 556 infants have LBW and mean birth weight is 2,983 grams. Of these infants whose birth weight is measured (1,922), 235 (12%) infants weighed less than 2,500 grams with mean birth weight of 3,030 grams. After combining estimated and reported birth weight, 791 (15%) infants have LBW and average birth weight is 3,000 grams.

Table 1 Description of birth weight; reported, imputed and combined

Category	Frequency (N)	Birth weight (< 2,500gm)	Birth weight (>= 2,500gm)	Mean
Birth weight measured	1,922	235	1,687	3,030
Imputed birth weight	3,318	556	2,762	2,983
Combined	5,240	791	4,449	3,000

Descriptive analysis

The descriptive analyses of variables are shown in table 2. A few mothers (17%) give birth in their teenage. More than half of infants (52%) are male babies. Nearly half of the mothers (47%) are uneducated. About 52% of the women have low socio- economic status. The majority of the mothers (79%) live in rural areas. Only a few of mothers (5%) are from Central Mountain.

Factors associated with birth weight

Table 2 describes the result from chi-square test. From the univariate analysis, it is found that factors like mother's age at child's birth, child's gender, wealth index, mother's education and sub-region are significantly associated with LBW as *p*-value is less than 0.05. Those variables which are significant in univariate analysis are consequently included in multiple logistic regression. In the first logistic model, all variables are initially included. From backward elimination procedure, only wealth index is found statistically insignificant ($p > 0.05$). Thus, in the final model, this variable is eliminated and the rest of all the factors are included. The model is then assessed using Pearson goodness of fit statistics and the result revealed *p*-value of 0.16. Table 3 illustrates strength of association between factors and birth weight. Mothers who give birth during their teenage (OR 1.466; 95% CI (1.261, 1.704)) have higher probability to give LBW infants than mothers who give birth after 20 years of age. Female babies (OR 1.277; 95% CI (1.181, 1.380)) are at a greater risk of having LBW than male babies. Uneducated mothers (OR 1.235; 95% CI (1.053, 1.448)) have a higher possibility to give LBW infants than educated mothers. Mothers living in Western mountain (OR 1.394; 95% CI (1.096, 1.774)), Far-western hill (OR 1.672; 95% CI (1.335, 2.093)) and Far-western terrain (OR 1.423; 95% CI (1.077, 1.880)) have a higher chance to deliver LBW infants than mothers living in other sub-regions.

Table 2 Chi-square test of birth weight with independent variables

Variable	Total number	Birth weight		p-value
	n (%)	>=2,500gm	<2,500gm	
		n	n	
Overall	**5,240**	**4,449**	**791**	
Mother's age at child's birth (Years)				
15-19	894 (17)	720	174	0.000
20-24	2035 (39)	1722	313	
25-29	1322 (25)	1157	165	
>=30	989 (19)	850	139	
Gender of baby				
Male	2723 (52)	2389	334	0.000
Female	2517 (48)	2060	457	
Mother's education				
No education	2442 (46)	2036	406	0.002
Primary education	1042 (20)	877	165	
Secondary education	1456 (28)	1269	187	
Higher education	300 (6)	267	33	
Wealth Index				
Poor	2729 (52)	2251	478	0.000
Middle	918 (18)	793	125	
Rich	1593 (30)	1405	188	
Residence				
Urban	1081 (21)	928	153	0.332
Rural	4159 (79)	3521	638	
Subregion				
Eastern mountain	319 (6)	282	37	0.000
Central mountain	242 (5)	195	47	
Western mountain	445 (8)	354	91	
Eastern hill	439 (8)	371	68	
Central hill	342 (7)	291	51	
Western hill	390 (7)	344	46	
Mid-Western hill	481 (9)	395	86	
Far-Western hill	463 (9)	357	106	
Eastern terrian	437 (8)	372	65	
Central terrian	529 (10)	513	16)	
Western terrian	374 (7)	333	41	
Mid-Western terrian	455 (9)	379	76	
Far-Western terrian	324 (6)	263	61	

Table 3 Factors associated with low birth weight

Variables	Unadjusted OR (95% CI)	Adjusted OR (95%CI)	*p*-value
Mother's age at child's birth (Years)			
15-19	1.35 (1.17, 1.56)	1.47 (1.26, 1.70)	0.000
20-24	1.02 (0.90, 1.14)	1.06 (0.94, 1.20)	0.319
25-29	0.80 (0.69, 0.91)	0.79 (0.69, 0.92)	0.002
>=30	0.91 (0.79, 1.06)	0.81 (0.69, 0.95)	0.009
Gender of baby			
Male	0.79 (0.74, 0.86)	0.78 (0.72, 0.85)	0.000
Female	1.26 (1.17, 1.36)	1.28 (1.18, 1.38)	0.000
Mother's education			
No education	1.23 (1.08, 1.41)	1.24 (1.05, 1.45)	0.009
Primary education	1.16 (0.99, 1.37)	1.08 (0.91, 1.28)	0.375
Secondary	0.91 (0.78, 1.06)	0.88 (0.75, 1.04)	0.140
Higher education	0.76 (0.58, 1.01)	0.85 (0.63, 1.14)	0.276
Subregion			
Eastern mountain	0.79 (0.57, 1.10)	0.80 (0.58, 1.12)	0.197
Central mountain	1.46 (1.07, 1.97)	1.43 (1.05, 1.96)	0.027
Western mountain	1.55 (1.24, 1.95)	1.39 (1.10, 1.77)	0.007
Eastern hill	1.11 (0.86, 1.42)	1.18 (0.91, 1.53)	0.208
Central hill	1.06 (0.80, 1.41)	1.09 (0.81, 1.47)	0.557
Western hill	0.81 (0.60, 1.09)	0.87 (0.65, 1.18)	0.369
Mid-Western hill	1.32 (1.04, 1.66)	1.28 (1.01, 1.62)	0.044
Far-Western hill	1.79 (1.44, 2.23)	1.67 (1.34, 2.09)	0.000
Eastern terrian	1.06 (0.82, 1.36)	1.11 (0.85, 1.45)	0.456
Central terrian	0.19 (0.12, 0.30)	0.18 (0.11, 0.29)	0.000
Western terrian	0.74 (0.55, 1.01)	0.79 (0.58, 1.10)	0.159
Mid-Western terrian	1.21 (0.95, 1.54)	1.17 (0.91, 1.49)	0.222
Far-Western terrian	1.40 (1.07, 1.83)	1.42 (1.08, 1.88)	0.013

Discussion

Studies that are conducted on LBW in a developing countries by using DHS used mother's opinion on baby's size as an alternative to birth weight (Khanal et al., 2014; Sreeramareddy et al., 2011) or they used only the subsample of reported birth weights (Khanal et al., 2014) because of highly missing birth weights in general. However, this study tries to handle this shortcoming by replacing missing values by means of regression imputation method. It does not mean that using imputation methods for replacing missing values removes the bias completely but it does lead bias downwards compared to a method of deleting missing observations completely (Lumley, 2011). After combining both reported and missing birth weight, the prevalence of LBW is found to 15%. Factors like mother's age at child's birth, mother's education, child's gender and sub-region are significantly associated with birth weight. Teenage mothers are more prone to deliver LBW infants. It can be interpreted from the table 3 that with the increase in mother's age, the likeliness of delivering LBW infants is reduced. This finding is supported by a meta-analyses study which showed that the degree of association of LBW is decreased with the increase in mother's age at child's birth (Gibbs et al., 2012). An inverse relationship between mother's education and LBW is found. Mothers who are uneducated are more likely to give birth to LBW infants than educated mothers. Several studies that are conducted on LBW also reported that the odds of delivering LBW infants are higher among the women who are uneducated compared to those who are educated (Mondal, 2000; Muula et al., 2011).

Similarly, the results of this study show that female infants are at a greater risk of having LBW than male infants. A similar findings is achieved in the study conducted in India (Mondal, 1998). The results of this study also show that mothers living in Western mountain, Far-western hill and Far-western terrain are more liable to give LBW because Far-western and Western region areas are considered as geologically and financially backward (Khanal et al., 2014).

Conclusion

This study finds prevalence of birth weight as 15 percent. Because 63 percent of infant's birth weight is missing in the data, regression imputation is used to impute missing values. The results from this study also show that, the mothers who give birth at the age of less than 20, who are illiterate, and who live in Western and Far-western regions have higher chance to deliver LBW infants. The finding of this study reflects that obtaining prevalence of LBW from the subsample of measured birth weight results into under estimation of prevalence. However, by using single imputation, missing values cannot be calculated with certainty. A further study is needed to confirm a level of uncertainty of the missing data.

Acknowledgements

We would like to express our sincere gratitude to Prof Don McNeil for providing guidance and support. We would also like to thank NDHS for granting us permission to conduct this study. Finally, we would also like to thank Graduate School for helping us financially in conducting this study.

References

Almond, D., Chay, K. Y., & Lee, D. S. (2005). The costs of low birth weight. *The Quaterly Journal of Economics*.

Almond, D., Chay, K. Y., Lee, D. S., & Berkeley, U. C. (2002). Does Low Birth Weight Matter ? Evidence from the U . S . Population of Twin Births. *Centre for Labor Economics, University of California, Berkeley.*

Balci, M. M., Acikel, S., & Akdemir, R. (2010). Low birth weight and increased cardiovascular risk: fetal programming. *International Journal of Cardiology*, *144*(1), 110-1.

Bethlehem, J. (2009). *Applied survey methods: A statistical perspective* (Vol. 558). John Wiley & Sons.

Blanc, A. K., & Wardlaw, T. (2005). Monitoring low birth weight : an evaluation of international estimates and an updated estimation procedure. *Bulletin of the World Health Organization*, *83*(3), 178-185.

Boerma, J. T., Weinstein, K. I., Rutstein, S. O., & Sommerfelt, A. E. (1996). Data on birth weight in developing countries: can surveys help? *Bulletin of the World Health Organization*, *74*(2), 209.

De Waal, T., Pannekoek, J., & Scholtus, S. (2011). *Handbook of statistical data editing and imputation* (Vol. 563). John Wiley & Sons.

Fan, Z. J., Lackland, D. T., Lipsitz, S. R., & Nicholas, J. (2006). The Association of Low Birthweight and Chronic Renal Failure Among Medicaid Young Adults with Diabetes and / or Hypertension. *Public Health Reports*, *121*(3), 239-244.

Gibbs, C. M., Wendt, A., Peters, S., & Hogue, C. J. (2012). The impact of early age at first childbirth on maternal and infant health. *Paediatric and Perinatal Epidemiology*, 26(s1), 259-284.

Kayode, G. A, Amoakoh-Coleman, M., Agyepong, I. A., Ansah, E., Grobbee, D. E., & Klipstein-Grobusch, K. (2014). Contextual risk factors for low birth weight: a multilevel analysis. *PloS One*, *9*(10), e109333.

Khanal, V., Sauer, K., Karkee, R., & Zhao, Y. (2014). Factors associated with small size at birth in Nepal: further analysis of Nepal Demographic and Health Survey 2011. *BMC Pregnancy and Childbirth*, 14(1), 32.

Khanal, V., Zhao, Y., & Sauer, K. (2014). Role of antenatal care and iron supplementation during pregnancy in preventing low birth weight in Nepal: comparison of national surveys 2006 and 2011. *Archives of Public Health*, 72(1), 4.

Lumley, T. (2011). *Complex surveys: A guide to analysis using R* (Vol. 565). John Wiley & Sons.

Ministry of Health and Population (MOHP) [Nepal], New Era, and ICF International Inc. (2011). Nepal Demographic and Health Survey. *Ministry of Health and Population, New Era, and ICF International, Caverton, Maaryland.*

Mondal, B. (1998). Low birth weight in relation to sex of baby, maternal age and parity: a hospital based study on Tangsa tribe from Arunachal Pradesh. *Journal of the Indian Medical Association*, *96*(12), 362-364.

Mondal, B. (2000). Risk factors for low birth weight in Nepali infants. *Indian Journal of Pediatrics*, *67*(7), 477-482.

Muula, A. S., Siziya, S., & Rudatsikira, E. (2011). Parity and maternal education are associated with low birth weight in Malawi. *African Health Sciences*, *11*(1), 65-71.

Nisar, Y. Bin, & Dibley, M. J. (2014). Antenatal iron-folic acid supplementation reduces risk of low birthweight in Pakistan: secondary analysis of Demographic and Health Survey 2006-2007. *Maternal & Child Nutrition*, 1-14.

Pedersen, M. G., Mortensen, P. B., & Webb, R. T. (2013). Birth weight, schizophrenia, and adult mental disorder: is risk confined to the smallest babies?. *Archives of General Psychiatry*, 67(9), 923-930.

Reichman, N. E., & Pagnini, D. L. (1997). Maternal Age and Birth Outcomes: Data from New Jersey. *Family Planning Perspectives*, *29*(6), 268.

Robles, A., & Goldman, N. (1999). Can accurate data on birthweight be obtained from health interview surveys? *International Journal of Epidemiology*, *28*(5), 925-931.

Silva, A. A. M., Bettiol, H., Barbieri, M. A., Pereira, M. M., Brito, L. G. O., Ribeiro, V. S., & Aragão, V. M. F. (2005). Why are the low birthweight rates in Brazil higher in richer than in poorer municipalities? Exploring the epidemiological paradox of low birthweight. *Paediatric and Perinatal Epidemiology*, *19*(1), 43–49.

Sreeramareddy, C. T., Shidhaye, R. R., & Sathiakumar, N. (2011). Association between biomass fuel use and maternal report of child size at birth-an analysis of 2005-06 India Demographic Health Survey data. *BMC Public Health*, *11*(1), 403.

United Nations Children's Fund and World Health Organisation. (2004). Low Birthweight: Country, regional and global estimates. *UNICEF and WHO, New York.*, 1-31.

Wilcox, A. J. (2001). On the importance-and the unimportance-of birthweight. *International Journal of Epidemiology*, 30(6), 1233-1241.

World Bank. (2015). Country and lending groups. Retrieved from http://data.worldbank.org/about/country-and-lending-groups.

World Health Organization. (2004). *International statistical classification of diseases and related health problems* (Tenth., Vol. 1). World Health Organization.

Vitae

Name: Miss Usha Singh

Student ID: 5720320002

Educational Attainment:

Degree	Name of institution	Year of Graduation
B.Sc. (Nursing)	Nepal Institute of Health Sciences, Nepal	2004

Award and Honor:

Thailand's Education Hub for Southern Region of ASEAN Countries (TEH-AC) Scholarship

International Conference:

Singh, U., Attachai, U., and Kuning, M. Estimating prevalence and factors associated with low birth weight in Nepal: further analysis of Nepal Demographic and Health Survey 2011. The 4th International Joint Conference on Society and Health. Global Health, Society and Human Development. 28-29 September, 2015. Mahidol University, Salaya Campus, Thailand.