Potential misclassification and bias due to low birth weight category in the population level

Hyojun Park, MA
Department of Population Health Sciences
University of Wisconsin-Madison

Abstract

A considerable misclassification and bias may be introduced in causal inference when low birth weight (LBW) category is used since LBW is a mixture of effects from the duration of the child's gestation and intrauterine growth. Small for gestational age (SGA) that captures only intrauterine growth is known as a better measure. Using the Early Childhood Longitudinal Study-Birth Cohort, this study identifies the degree of misclassification of LBW model and determines the potential bias in various models that LBW was used as a dependent, independent, or mediating variable. The results show that the magnitude of misclassification was considerable and differed by preterm and term birth. In addition, the magnitude and direction of bias greatly varied in the LBW models as a dependent, independent, or mediating variable. The results strongly suggest that LBW may introduce bias in the causal inference, so that SGA category should be utilized whenever possible.

Abbreviations:

LBW: low birth weight; SGA: small for gestational age; GA: gestational age; BMI: body mass index; PPV: positive predictive value; NPV: negative predictive value (NPV)

Introduction

Low birth weight (LBW; birth weight is under 2,500g) is not only a marker that indicates adverse birth outcomes of infants, but is also a precursor to significant health consequences including perinatal mortality and morbidity[1]. It is well known that low birth weight is a limited measure that is affected by the duration of the child's gestation and/or intrauterine growth.[2] For this reason, many clinical and epidemiological studies have utilized small weight for gestational age (SGA; birth weight is under 10th percentile for the gestational age), which separates intrauterine growth from the duration of gestation, instead of LBW in order to better capture the underlying pathology of infant's growth.[3]

Thus, a considerable misclassification and bias may be introduced in causal inference when LBW category is used. The degree of potential bias may vary by the model which role LBW plays in the pathways.[4] The bias may also vary by the underlying pathology which health outcomes are of concern. To our best knowledge, however, few studies have quantified the misclassification arising from using LBW versus SGA, and what direction would be. The assessment of the direction and magnitude of potential bias due to misclassification is crucial to both validate findings from the existing literature and improve our understanding of the relationship between intrauterine growth and health outcomes.

The objectives of this study are to: 1) identify if and to what extent LBW and SGA differently classify infants; and 2) determine the potential bias due to the misclassification in various models. More specifically, this study will assess the magnitude of misclassification, the degree of agreement and validity between LBW and SGA, and then the potential bias due to the misclassification.

Methods

Data and measures

Data were from the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) of 2001. The ECLS-B is a nationally representative longitudinal dataset providing detailed information on approximately 10,700 children's early life experiences at birth, 9-months, 2-years, preschool, and kindergarten entry. Among 10,700 infants, we excluded twins or multiples, infants with clinical gestational ages less than 23 weeks or greater than 41 weeks, and infants who were assigned with zero or negative weight values in the ECLS-B dataset, leaving 7,000 infants eligible for our study.

LBW status was defined as birth weight less than 2,500g (LBW) and over 2,500g (non-LBW). For the calculation of Kappa statistic, infants were classified as LBW(less than 2,500g), normal birth weight (NBW; 2,500-4,000g), and high birth weight (HBW; over 4,000g). Gestational age (GA) was defined as clinical gestational age at birth provided by ECLS-B. Infants were classified as preterm (less than 36 weeks) and term (37 weeks and over). SGA status was defined as infants who were under the 10^{th} percentile birth weights (SGA) and over 10^{th} percentile birth weights (non-SGA) per gestational age given the intrauterine growth curves provided by Olsen et al.[5] For the calculation of Kappa statistic, infants were classified as SGA(less than 10^{th} percentile), appropriate for gestational age (AGA; 10^{th}-90^{th} percentile), and larger for gestational age (LGA; over 90^{th} percentile).

Self-reported tobacco use before pregnancy was measured as a dichotomous scale. Body mass index (BMI) percentile was measured as the most recent BMI percentile available in the data. In addition, monthly age at the most recent assessment was included in the model to control for the age effect on the BMI percentile.

The following measures were used as control measures including children's sex, race/ethnicity, plurality, delivery methods, medical risk factors for pregnancy, obstetric procedures, complications of labor and/or delivery, abnormal conditions of the newborn, congenital anomalies of the child, adequacy of prenatal care utilization; maternal age, weight gain, estimated BMI, education, marital status, birth place, tobacco and alcohol use before pregnancy, and household income at 9 months.

Statistical analyses

The different classifications between LBW and SGA were presented as gender-specific scatter plots by birth weight and gestational age at birth. The agreement of classifications between LBW and SGA were measured by kappa, sensitivity/specificity, and positive/negative predictive values. These statistics were separately calculated by preterm and term birth. Kappa is the fraction of the observed agreement not due to chance in relation to the maximum non-chance agreement when using a categorical classification of a variable.[6] Sensitivity is the probability that someone exposed is classified as exposed, whereas specificity

is the probability that someone unexposed is classified as unexposed.[4] Positive predictive value (PPV) is the proportion of true positives among individuals who test positive, whereas negative predictive value (NPV) is the proportion of true negatives out of the total who test negative.[6]

The impacts of misclassification are different by the roles of measure[4]. In the causal pathway, a measure may play a role as a dependent, mediating, moderating, or independent variable. To determine the potential impact of misclassification, three models were examined; LBW/SGA as a dependent, mediating, and independent variable. The comparisons of estimates between LBW and SGA were made under the assumption that SGA is a gold standard. Logistic regression was used to investigate the impact of misclassifications when LBW is used as dependent and mediating variable, whereas linear regression model was used when LBW is used as an independent variable. The direction and magnitude of the impact was measured as odds ratios or regression coefficients after adjusting for other risk factors. Statistical significance was determined as the 95% confidence interval, and clinical significance was determined as the 10% changes of estimates.

Figure 2. The different role of SGA and LBW on the pathways

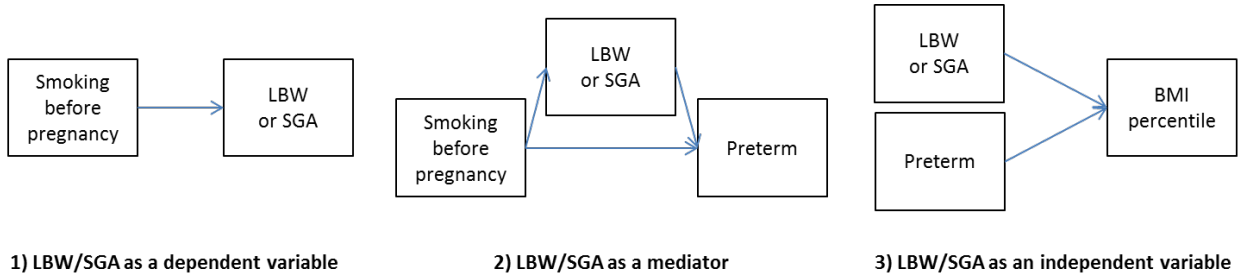


Figure 2 demonstrates the pathways of each model. In the first model, LBW or SGA is a dependent variable that is affected by smoking before pregnancy (Figure 2-1, Table 2). In the second model, LBW or SGA is a mediator in the pathways between smoking before pregnancy and preterm birth (Figure 2-2, Table 3). In the third model, LBW or SGA is an independent variable that influences BMI percentile later (Figure 2-3, Table 3).

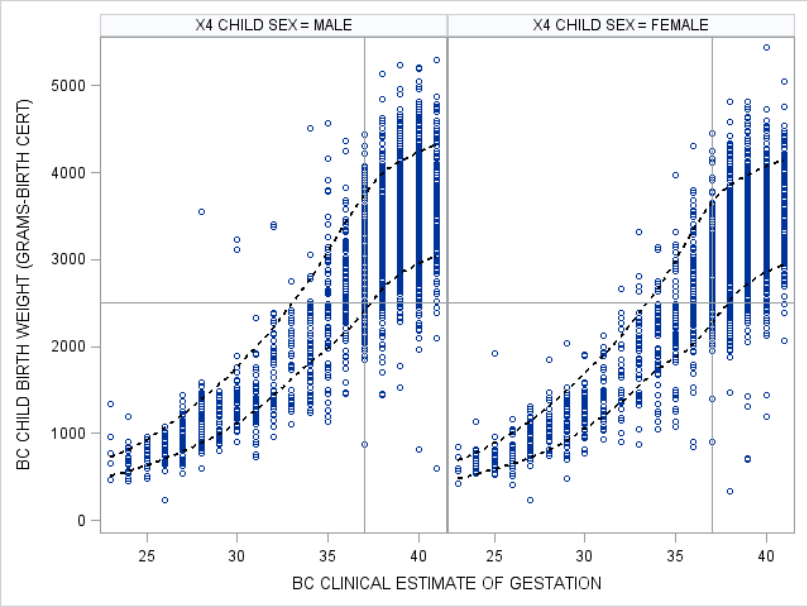
Several modifications were conducted to validate the current study. Models with alternative definitions of risk factors were compared; for example, SGA can be measured as below -2 standard deviations, 3rd or 5th percentile, and Wilcox's method[7-9]. Also potential biases due to sampling will be examined by using subgroup analyses changing exclusion criteria.

Analyses were conducted using survey procedures available in SAS v9.3.The analytic scheme was a sub-domain analyses that analyzes the included cases and excluded cases as different domains to retain appropriate weight scheme in ECLS-B. Missing values in measures were treated as an additional category of 'missing' to retain the cases as many as possible. Complex sampling weight scheme in ECLS-B was applied then adjusted to the total sample size. The numbers of cases were rounded to the nearest 50 as the guideline of National Center for Education Statistics (NCES).

Results

The potential misclassification when using LBW category was presented as scatterplots by sex in Figure 1. The Y axis is the birth weight of infants and the X axis represents gestational age. The horizontal line at birth weight 2,500g separates LBW from normal/high birth weight infants. The vertical line at 37 weeks of gestational age separates preterm from term birth infants. The upper dotted curve indicates the 90th percentile birth weight among infants at the same gestational age, whereas the lower one indicates the 10th percentile. Infants under the 10th percentile curve are generally considered as SGA infants.

Figure 1. Classification of infants using birth weight and gestational age



The figure clearly demonstrates a considerable amount of misclassification between SGA and LBW categories. LBW category seems to poorly capture SGA infants, even when it uses with preterm/term category. More specifically, infants who were between a horizontal line at 2,500g and the 10th percentile curve are all misclassified. Among preterm infants, those are in the left side of vertical line, a great amount of non-SGA infants were classified as LBW, whereas among term infants those are in the right side of vertical line, also a great amount of SGA infants was not classified as LBW.

Table 1. The agreement and validity statistics of LBW and SGA categories, national estimates from the ECLS-B

	Preterm			Term		
	Statistics	95% LCL	95% UCL	Statistics	95% LCL	95% UCL
Kappa	0.08	0.06	0.09	0.64	0.62	0.67
Sensitivity	1.00	1.00	1.00	0.48	0.43	0.52
Specificity	0.80	0.78	0.82	0.99	0.98	0.99
Positive predictive values	0.20	0.18	0.22	0.77	0.72	0.81
Negative predictive values	1.00	1.00	1.00	0.95	0.94	0.96

Abbreviations: LBW (low birth weight); SGA (small for gestational age); 95%LCL (95% lower confidence limit); 95%UCL(95% upper confidence limit).

To quantify the potential misclassification when using LBW category, several agreement and validity statistics were calculated and presented in Table 1. The statistics were separately calculated by preterm and term birth since the degree and direction of misclassification clearly differ by each other as shown in Figure 1.

Agreement and validity statistics demonstrate a considerable amount of misclassification between SGA and LBW categories. The LBW measure exhibits higher sensitivity but lower specificity among preterm infants, whereas higher specificity but lower sensitivity among term infants. This result implies that the magnitude of bias would be very severe and the direction of bias is dependent to preterm and term birth.

More specifically, for the infants born preterm, the agreement between LBW and SGA categories was very poor (Kappa=0.08; CI: 0.06-0.09). LBW category correctly detected all SGA infants as LBW (sensitivity=1.00; CI: 1.00-1.00), but 80% of LBW infants were in fact non-SGA infants (e.g. 0.8=1.0-0.2; PPV=0.20; CI: 0.18-0.22). In the same sense, LBW category correctly detected 80% of non-SGA infant as non-LBW (specificity=0.80; CI: 0.78-0.82), and all non-LBW infants were non-SGA (NPV=1.00; CI: 1.00-1.00). In the same sense, for the infants born term, the agreement between LBW and SGA categories was moderate (Kappa=0.64; CI: 0.62-0.67). LBW category correctly detected only 48% of SGA infants as LBW (sensitivity=0.48; CI: 0.43-0.52), but 23% of LBW infants were in fact non-SGA infants (e.g. 0.23=1.0-0.77; PPV=0.77; CI: 0.72-0.81). In the same sense, LBW category correctly detected 99% of non-SGA infants as non-LBW (specificity= 0.99; CI: 0.98-0.99), and 5% of non-LBW infants were in fact SGA infants (e.g. 0.05=1.00-0.95; NPV=0.95; CI: 0.94-0.96).

Table 2. The potential impact of misclassification in the dependent variable

	Dependent variable= LBW			Dependent variable= SGA		
	AOR	95%LCL	95%UCL	AOR	95%LCL	95%UCL
Smoking before pregnancy	2.72	1.67	4.44	1.74	1.03	2.95
Preterm birth	6.27	5.49	7.17	0.88	0.77	1.01

Notes:

1) Dependent variable: LBW or SGA; Independent variables: Smoking before pregnancy and Preterm birth.
2) All models were adjusted for: children's sex, race/ethnicity, plurality, delivery methods, medical risk factors for pregnancy, obstetric procedures, complications of labor and/or delivery, abnormal conditions of the newborn, congenital anomalies of the child, adequacy of prenatal care utilization; maternal age, weight gain, estimated BMI, education, marital status, birth place, tobacco and alcohol use before pregnancy, and household income at 9 months.

Table 2 summarizes the results of logistic regression that LBW or SGA is regressed on smoking before pregnancy and preterm birth. Smoking before pregnancy is a well-known risk factor that increases both LBW/SGA and preterm birth.[2,3] The results indicate that: 1) the impact of smoking before pregnancy may be overestimated when LBW was used as a dependent variable, and 2) this overestimate may be due to misclassification creates a spurious pathway between preterm birth and LBW. More specifically, when using LBW, infants whose mothers were smoking before pregnancy were 2.72 times (CI: 1.67-4.44) more likely born LBW than infants whose mothers were not smoking. However, when using SGA, the risk was estimated as 1.74 (CI: 1.03-2.95). Interestingly, the risk of preterm infants to be LBW was 6.27 (CI: 5.49-7.17), whereas that of SGA was 0.88 (CI: 0.77-1.01) and not statistically significant.

Table 3. The potential impact of misclassification in the mediating variable

	Baseline model			Mediator: LBW			Mediator: SGA		
	AOR	95% LCL	95% UCL	AOR	95% LCL	95% UCL	AOR	95% LCL	95% UCL
Smoking before pregnancy	1.25	0.88	1.78	0.85	0.63	1.14	1.26	0.88	1.80
LBW				6.30	5.55	7.16			
SGA							0.89	0.78	1.01

Notes:

1) Dependent variable: Preterm birth; Independent variable: Smoking before pregnancy; Mediator: LBW or SGA.
2) All models were adjusted for: children's sex, race/ethnicity, plurality, delivery methods, medical risk factors for pregnancy, obstetric procedures, complications of labor and/or delivery, abnormal conditions of the newborn, congenital anomalies of the child, adequacy of prenatal care utilization; maternal age, weight gain, estimated BMI, education, marital status, birth place, tobacco and alcohol use before pregnancy, and household income at 9 months.

Table 3 summarizes the impact of misclassification when LBW or SGA is used as mediator. The impact of smoking before pregnancy was estimated as opposite direction and the impact of LBW was inflated when LBW was used as a mediator between smoking before pregnancy and preterm birth. More specifically, without controlling for LBW or SGA, the infants whose mothers smoked before pregnancy were 1.25 times more likely born preterm (95% CI: 0.88-1.78); however, the impact of smoking before pregnancy changed as 0.85 when LBW was used as a mediator, which is the opposite direction. When SGA was used as a mediator, however, the impact of smoking before pregnancy does not change (AOR=1.26; CI: 0.88-0.95), which is the expected result that is consistent with other research (ref). The change of direction of the association may due to the misclassification that opens a spurious pathway between preterm birth and LBW (LBW: 6.30 (CI: 5.55-7.16) vs. SGA: 0.89 (CI: 0.78-1.01)).

Table 4. The potential impact of misclassification in the independent variable

	Independent=LBW			Independent=SGA		
	Est	s.e.	p	Est	s.e.	p
LBW	-12.21	1.54	0.00			
SGA				-10.83	1.60	0.00
Preterm birth	1.04	1.63	0.53	-4.87	1.33	0.00
Age at the assessment	0.13	0.03	0.00	0.13	0.03	0.00

Notes:

1) Dependent variable: the most recent BMI percentile; Independent variables: LBW/SGA and Preterm
2) All models were adjusted for: children's sex, race/ethnicity, plurality, delivery methods, medical risk factors for pregnancy, obstetric procedures, complications of labor and/or delivery, abnormal conditions of the newborn, congenital anomalies of the child, adequacy of prenatal care utilization; maternal age, weight gain, estimated BMI, education, marital status, birth place, tobacco and alcohol use before pregnancy, household income at 9 months, and age at the assessment.

Table 4 summarizes the impact of misclassification when LBW or SGA is used as an independent variable. The magnitude of the impacts of LBW and SGA were similar with each other and the direction

of association is consistent with previous studies. However, the impacts of preterm birth on the BMI percentile were considerably different between two models. The results show that the impact of LBW on BMI may not be biased, but the impact of the preterm may be biased. More specifically, the BMI percentile was regressed on LBW or SGA and preterm birth. The result is that on average the BMI percentile of LBW infants was -12% lower than that of non-LBW infants. In the same sense, the BMI percentile of SGA infants was -11% lower than that of non-SGA infants. However, The impact of preterm birth was 1.04 (s.e.=1.63; p=0.53) and not significant in the LBW model, whereas the impact of preterm birth on the BMI percentile was -4.87 (s.e.=1.33; p=0.000) in the SGA model.

Discussion

We found considerable misclassification arising from the use of LBW categories instead of SGA. As expected, LBW was greatly associated with preterm birth, so that the magnitude and direction of misclassification differed by preterm and term birth. The impact of misclassification when using LBW category also differed by the models which role LBW plays in the pathways. In the example of this study, 1) the estimates of smoking before pregnancy and preterm on LBW were biased away the null; 2) in the mediation model, the estimate of smoking before pregnancy was qualitatively biased and the estimate of LBW was biased away the null; and 3) the estimates of LBW on BMI was not biased, but the impact of preterm was bias toward the null or qualitatively biased.

Limitation should be mentioned. Most of all, SGA may not be a perfect measure for intrauterine growth. Also the gestational age may not be accurate. Moreover, some omitted measures or pathways may exist that greatly influence each of three models. If this is the case, the impacts of SGA in different models that were assumed as a gold standard may be biased. Lastly, the direction and magnitude of bias that were estimated in this study may not be generalizable in other population.

Conclusions

This study is one of few studies that identifies the magnitude and direction of misclassification when using LBW rather than SGA categories and determines the direction and magnitude of potential bias due to the misclassification in the various roles of LBW within the pathways. The results of the study strongly suggest that LBW may introduce bias in the causal inference, so that SGA category should be utilized whenever possible.

References

1. Goldenberg RL, Culhane JF. Low birth weight in the United States. *Am J Clin Nutr.* Feb 2007;85(2):584S-590S.

2. Kramer MS. Intrauterine growth and gestational duration determinants. *Pediatrics.* Oct 1987;80(4):502-511.

3. Saenger P, Czernichow P, Hughes I, Reiter EO. Small for gestational age: short stature and beyond. *Endocr Rev.* Apr 2007;28(2):219-251.

4. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology.* 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.

5. Olsen IE, Groveman SA, Lawson ML, Clark RH, Zemel BS. New intrauterine growth curves based on United States data. *Pediatrics.* Feb 2010;125(2):e214-224.

6. Szklo M, Nieto FJ. *Epidemiology : beyond the basics.* 2nd ed. Sudbury, Mass.: Jones and Bartlett Publishers; 2007.

7. Wilcox AJ, Russell IT. Birthweight and perinatal mortality: I. On the frequency distribution of birthweight. *Int J Epidemiol.* Sep 1983;12(3):314-318.

8. Wilcox AJ, Russell IT. Birthweight and perinatal mortality: II. On weight-specific mortality. *Int J Epidemiol.* Sep 1983;12(3):319-325.

9. Wilcox AJ, Russell IT. Birthweight and perinatal mortality: III. Towards a new method of analysis. *Int J Epidemiol.* Jun 1986;15(2):188-196.