

**FHS PUBLIC ACCESS**

Author manuscript

Matern Child Health J. Author manuscript; available in PMC 2017 July 01.

Published in final edited form as:

Matern Child Health J. 2016 July ; 20(7): 1497–1505. doi:10.1007/s10995-016-1948-6.

Impact of Missing Data for Body Mass Index in an Epidemiologic Study

Hilda Razzaghi¹, Sarah C. Tinker¹, Amy H. Herring², Penelope P. Howards³, D. Kim Waller⁴, and Candice Y. Johnson^{1,3,5} **the National Birth Defects Prevention Study**

Hilda Razzaghi: hir2@cdc.gov

¹National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention (CDC), Mail-Stop F76, 1600 Clifton Road, Atlanta, GA 30333, USA

²Gillings School of Global Public Health and Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

³Rollins School of Public Health, Emory University, Atlanta, GA, USA

⁴UTHealth, School of Public Health, Houston, TX, USA

⁵National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, OH, USA

Abstract

Objective—To assess the potential impact of missing data on body mass index (BMI) on the association between prepregnancy obesity and specific birth defects.

Methods—Data from the National Birth Defects Prevention Study (NBDPS) were analyzed. We assessed the factors associated with missing BMI data among mothers of infants without birth defects. Four analytic methods were then used to assess the impact of missing BMI data on the association between maternal prepregnancy obesity and three birth defects; spina bifida, gastroschisis, and cleft lip with/without cleft palate. The analytic methods were: (1) complete case analysis; (2) assignment of missing values to either obese or normal BMI; (3) multiple imputation; and (4) probabilistic sensitivity analysis. Logistic regression was used to estimate crude and adjusted odds ratios (aOR) and 95 % confidence intervals (CI).

Results—Of NBDPS control mothers 4.6 % were missing BMI data, and most of the missing values were attributable to missing height (~90 %). Missing BMI data was associated with birth outside of the US (aOR 8.6; 95 % CI 5.5, 13.4), interview in Spanish (aOR 2.4; 95 % CI 1.8, 3.2), Hispanic ethnicity (aOR 2.0; 95 % CI 1.2, 3.4), and <12 years education (aOR 2.3; 95 % CI 1.7,

Correspondence to: Hilda Razzaghi, hir2@cdc.gov.

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Electronic supplementary material The online version of this article (doi:10.1007/s10995-016-1948-6) contains supplementary material, which is available to authorized users.

Conflicts of interest

The authors have no conflicts of interest to report. No financial disclosures were reported by the authors of this paper.

3.1). Overall the results of the multiple imputation and probabilistic sensitivity analysis were similar to the complete case analysis.

Conclusions—Although in some scenarios missing BMI data can bias the magnitude of association, it does not appear likely to have impacted conclusions from a traditional complete case analysis of these data.

Keywords

Missing BMI; NBDPS; BMI; Missing data; Birth defect

Introduction

Obesity is a national epidemic in the U.S. with several recent large population-based studies indicating a steady rise in its prevalence [1–7]. This increase has also been observed among women of childbearing age, which is of great public health concern given the increased risk of adverse pregnancy outcomes, including miscarriage, stillbirth and certain birth defects associated with prepregnancy obesity [8–12].

Body mass index (BMI) is a common metric used to assess an individual's body fat; it is calculated as weight (kg)/height² (m) [13–15]. BMI cannot be calculated when information on either height or weight is missing. Study participants with missing data are often excluded from statistical analyses, which may introduce bias if the likelihood of missing data is associated with both the exposure and the outcome [16]. Even in the absence of bias, if missing data are more common in some subgroups of the population than others, this differential exclusion may affect generalizability of results.

While the issues surrounding inaccurate self-report of height and weight (and its derivative, BMI) have been widely studied [17], the potential impact of missing height and weight values on the results of analyses assessing health outcomes associated with BMI has received less attention. To our knowledge, few studies of characteristics of study participants with missing BMI are from adult U.S. populations. In a study of Portuguese adult women, missing data on BMI was associated with age, education, smoking, and physical activity level [18]. In adolescents in the U.S., Germany, and Portugal, missing BMI was associated with age, body image, health behaviors, and composition of social networks [19–21]. Correlations between missing BMI and other sociodemographic and health-related variables support the potential for bias when participants with missing BMI are excluded from analyses.

The objectives of this study were: (1) to describe the frequency of missing prepregnancy height, weight, and BMI among control mothers in the National Birth Defects Prevention Study (NBDPS); (2) to assess characteristics associated with missing BMI; and (3) to assess the potential impact of missing BMI on the association between prepregnancy obesity and specific birth defects.

Methods

National Birth Defects Prevention Study

The NBDPS includes cases of specific structural birth defects identified by active, population-based birth defects surveillance systems at ten centers located throughout the U.S. (entire state: Arkansas, Iowa, New Jersey, Utah; specific counties: California, Georgia, Massachusetts, New York, North Carolina, Texas). Controls are live-born infants without major birth defects who were randomly selected either using birth hospital records or birth certificates [22, 23]. Data included in this analysis were obtained from infants born on or after October 1, 1997 and with expected dates of delivery (EDD) on or before December 31, 2009. Cases in the NBDPS include live-born infants from all study sites, stillbirths ≥ 20 weeks gestation from all sites except for NJ (all years) and NY (before the year 2000), and elective terminations ≥ 20 weeks of gestation from all sites except for MA and NJ (all years), and NY (before the year 2000). Additionally, cases with major chromosomal abnormalities or single-gene disorders are excluded from NBDPS [24]. The NBDPS was approved by the institutional review boards of the Centers for Disease Control and Prevention (CDC) and the participating study sites. Maternal interviews were conducted using a standardized computer-assisted telephone interview in English or Spanish between 6 weeks to 24 months after delivery. As a part of the maternal interview, mothers were asked to report their height and prepregnancy weight. For the first several years of the study (through June 2002), mothers were only able to systematically report their height in feet and inches; in the later years of data collection they were given the option to report their height in centimeters.

Body Mass Index

We defined BMI according to the National Heart, Lung, and Blood Institute and the World Health Organization (WHO) criteria [25, 26]. Underweight was defined as a BMI of less than 18.5 kg/m^2 , normal weight was defined as a BMI of $18.5\text{--}24.9 \text{ kg/m}^2$, overweight was defined as a BMI of $25.0\text{--}29.9 \text{ kg/m}^2$, and obesity as a BMI of 30 kg/m^2 or higher.

We assessed the frequency and factors associated with missing BMI information among NBDPS control mothers who completed the interview. Our analysis focused on BMI rather than height or weight alone since previous studies of birth defects examined associations with BMI. We excluded mothers with implausible BMI values—those less than 10.0 or greater than 70.0 ($N = 5$). We used logistic regression to estimate odds ratios (ORs) for the associations between control mothers having missing BMI information and selected maternal characteristics: age, race/ethnicity, education, language of interview, country of birth, parity, folic acid use in the month prior to conception, cigarette smoking in the month prior to conception, or alcohol consumption in the month prior to conception. These factors were considered as covariates in a previous analysis of NBDPS data in which the association between BMI and specific birth defects was assessed [12].

We compared the missing BMI frequencies observed in NBDPS to those observed in the National Health and Nutrition Examination Survey (NHANES), and specifically the distribution of missing self-reported height and weight by race/ethnicity. NHANES is a stratified, multistage probability sample survey of the civilian, noninstitutionalized

population of the U.S. conducted by the National Center for Health Statistics [27]. NHANES includes an in-home questionnaire, in which participants are asked to report their height and weight, and an examination at a Mobile Examination Center (MEC), where height and weight are directly measured. We included all non-pregnant women 16–53 years of age who completed both the interview and examination portions of the 1999–2010 NHANES; this age range was selected because NHANES collected self-reported height and weight in all participants 16 and older, and the maximum maternal age in NBDPS was 53 years old. We used the MEC sample weights and the appropriate sample design variables to account for the complex survey design, oversampling, and differential nonresponse and noncoverage in the NHANES sample [28].

Analysis

We estimated the association between maternal prepregnancy obesity and spina bifida, gastroschisis, and cleft lip with or without cleft palate (CLP) using four different methods to investigate the potential impact of missing BMI in the NBDPS. We chose to consider spina bifida and gastroschisis because these defects were strongly associated with maternal prepregnancy obesity in a previous NBDPS analysis; compared to normal weight mothers, spina bifida risk was higher among obese mothers and gastroschisis risk was much lower. We wanted to determine whether selection bias due to complete case analysis (excluding mothers with missing data on prepregnancy BMI) could have induced the observed associations. CLP was the largest defect category for which an association with BMI was not observed in the previous NBDPS analysis, and we also sought to assess whether bias due to missing data could have obscured an association with this defect. For all analyses normal BMI was used as the referent category.

We first analyzed the data using a complete case analysis approach, the commonly used method of excluding observations with missing BMI. We used logistic regression to estimate crude and adjusted odds ratios; the covariates listed above were those included in the adjusted models. For the second method, we estimated both adjusted and crude odds ratios resulting from all possible combinations of assignments of missing values for BMI to either obese or normal prepregnancy BMI (SAS code available upon request). For example, there were 76 spina bifida cases and 435 controls for which maternal prepregnancy BMI was missing; we estimated odds ratios for the scenarios that result from every possible combination of assigning each of these missing values to either obese or normal weight (from 0 to 76 cases being classified as obese and from 0 to 435 controls being classified as obese, resulting in 33,572 odds ratios (77×436) estimated for spina bifida). We plotted the percentage of cases classified as obese against the percentage of controls classified as obese to summarize and visualize the odds ratios resulting from these scenarios.

The third method we employed was multiple imputation for the missing continuous height and weight variables, which were then used to calculate BMI (categorized after imputations) utilizing the PROC MI and PROC MIANALYZE procedures in SAS based on a joint multivariate normal distribution. This procedure used the values of the non-missing height and weight variables as well as age, race/ethnicity, and education to impute the value for the

missing height and weight data and integrated the additional variance from this process into the final estimate [29].

As the final method, probabilistic sensitivity analysis was used to examine the effects of missing BMI on study results. We simulated 1000 datasets in which mothers with missing BMI data in NBDPS were assigned to a BMI category, based on their probability of being in each category according to the NHANES prevalence estimated for their stratum of age, race/ethnicity, and education. To estimate these prevalences we used the height and weight data from the direct measurements taken at the NHANES MEC. We estimated the prevalence of each BMI category for strata defined by age (<18, 18–24, 25–29, 30–34, 35 years), race and ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, and other), and education level (less than high school, high school graduate or equivalent, some college, college graduate or higher), accounting for sampling weights and the complex survey design. For each simulated dataset each mother with missing BMI information was assigned a value between 0 and 1 based on a uniform distribution. For each strata defined by age, race/ethnicity, and education, BMI categories were defined based on their cumulative probabilities. For example, for a woman aged <18 years, of non-Hispanic white race/ethnicity, with less than high school education the cumulative probability for each BMI category was: underweight—0.16; normal weight— $0.16 + 0.61 = 0.77$; overweight— $0.77 + 0.14 = 0.91$; and obese— $0.91 + 0.09 = 1$. A mother in this stratum randomly assigned a value of 0.49 from the standard uniform distribution would be assigned to the normal BMI category (bounded by 0.17–0.77). After assigning each mother with missing BMI to a BMI category, we estimated the crude and adjusted odds ratios for the association between obesity and spina bifida, gastroschisis, and CLP. The 1000 datasets produced a distribution of odds ratio point estimates, as well as lower and upper bounds for the 95 % confidence intervals for those estimates. We summarized the simulation results using the median of the point estimates and defined our uncertainty interval as the union of the confidence intervals across simulations (i.e., the 2.5th percentile of the lower confidence interval distribution and the 97.5th percentile of the upper confidence interval distribution), based on the idea of the “region of uncertainty” as described by Molenberghs and Kenward [30].

All analyses of NBDPS data were conducted using SAS (version 9.3; SAS Institute, Cary, NC); all analyses of NHANES data were conducted using SAS-callable SUDAAN (version 11; Research Triangle Institute; Research Triangle Park, NC).

Results

Overall, 435 (4.6 %) of the 10,075 control mothers from NBDPS included in our analysis were missing data on BMI (Table 1). The maternal factors with the strongest independent associations with missing BMI data were Hispanic ethnicity [adjusted odds ratio (aOR) 2.0; 95 % confidence interval (CI) 1.2, 3.4]; maternal education less than high school (aOR 2.3; 95 % CI 1.7, 3.1); having completed the interview in Spanish (aOR 2.4; 95 % CI 1.8, 3.2); and maternal birth outside of the U.S. (aOR 8.6; 95 % CI 5.5, 13.4). Mothers who reported folic acid supplement use or alcohol use in the month before pregnancy were less likely to have a missing BMI value (aOR 0.5; 95 % CI 0.3, 0.7 and aOR 0.7; 95 % CI 0.4, 1.0, respectively), as were non-Hispanic black mothers (aOR 0.3; 95 % CI 0.1, 0.9). Additional

analysis of factors associated with missing height and weight data (presented in Supplementary Tables 1 and 2) showed that the results for height analysis are similar to those of BMI, and results for weight analysis were slightly attenuated for language of interview and country of birth compared with those for BMI.

The majority of missing BMI values in the NBDPS were attributed to missing height information (Table 2). Among the 435 control mothers with missing BMI data, 392 (90.1 %) were missing information on height; only 86 (19.8 %) were missing information on weight. Among the NHANES participants, self-reported BMI was missing approximately half as often (2.4 %) as among control mothers in the NBDPS. Unlike the NBDPS data, in the NHANES data, missing weight information was slightly more common than missing height information (1.6 and 1.1 %, respectively). Overall, Hispanic women were much more likely to have missing BMI information than women in any other racial/ethnic category for both NBDPS and NHANES. Almost 16 % of Hispanic NBDPS participants, and almost 10 % of non-pregnant Hispanic women age 16–53 years in NHANES were missing information on BMI (Table 2).

In the complete case analysis of the association between prepregnancy obesity and spina bifida, in which participants with missing BMI were excluded, we observed an aOR of 1.6 (95 % CI 1.4, 1.9) (Table 3). For the analysis that assessed all possible datasets using different combinations of assignment of missing values to obese or normal BMI, for the most extreme scenarios the “true” adjusted odds ratios could be between 1.1 (95 % CI 0.9, 1.3; all missing cases normal weight; all missing controls obese) and 2.3 (95 % CI 2.0, 2.7; all missing cases obese; all missing controls normal weight). It is, therefore, possible that the observed association between prepregnancy obesity and spina bifida is entirely attributable to missing data, but only under extreme conditions. It is also possible that the missing data resulted in an underestimate of the true odds ratio by as much as 30 %. In addition, when examining crude associations, all possible crude odds ratios were greater than one and statistically significant (Fig. 1a). The multiple imputation and probabilistic sensitivity methods each produced results nearly identical to the complete case analysis (aOR 1.6; 95 % CI 1.4 1.9; and aOR 1.6; 95 % CI 1.3, 1.9, respectively).

For gastroschisis, the complete case analysis yielded an aOR of 0.2 (95 % CI 0.2, 0.3) (Table 3). Under the most extreme scenarios the “true” adjusted odds ratio could be between aOR 0.2; 95 % CI 0.1, 0.2 and aOR 0.4; 95 % CI 0.3, 0.5; the results demonstrate that the negative association of prepregnancy obesity with gastroschisis cannot be due to bias caused by missing BMI data (Fig. 1b). The multiple imputation and probabilistic sensitivity methods each produced results nearly identical to those from the analysis in which the missing data were excluded (aOR 0.2; 95 % CI 0.2, 0.3 and aOR 0.3; 95 % CI 0.2, 0.4, respectively).

The aOR estimate from the complete case analysis for CLP was 1.1 (95 % CI 1.0, 1.3) (Table 3). Under the most extreme scenarios the “true” adjusted odds ratios could be between aOR 0.8 (95 % CI 0.7, 0.9) and aOR 1.6 (95 % CI 1.4, 1.8); therefore, the data could be consistent with maternal prepregnancy obesity being associated with a decreased or increased risk of CLP, depending on the distribution of the missing BMI data. When all

possible distributions of missing BMI were considered (Fig. 1c), approximately half of the crude odds ratios were greater than one and statistically significant. There were, however, many possible crude odds ratios consistent with the null association; only a small number (4.5 %) were consistent with a significant protective association. The multiple imputation and probabilistic sensitivity methods each produced results nearly identical to those from the analysis in which the missing data were excluded (aOR 1.1; 95 % CI 1.0, 1.3 and aOR 1.1; 95 % CI 1.0, 1.3, respectively).

Discussion

We found that the majority of missing BMI data among controls was attributable to missing height in NBDPS. Furthermore, missing information on maternal prepregnancy BMI was associated with maternal race/ethnicity, education and lower acculturation (Spanish interview and non-U.S. country of birth). In the NBDPS data, 89 % (347/392) of missing height data in control mothers was among Hispanic women, primarily from interviews administered in Spanish (data not shown). Thus, the factors that remained strongly associated with missing BMI even after controlling for Hispanic ethnicity suggest that lower acculturation, or the degree of adaptation to a new culture, is related to increased likelihood of missing data for BMI (Table 1). Anecdotally, NBDPS interviewers reported that women who did not report their height tended not to know it, rather than having refused to report it. This suggests the possibility of substantially reducing the prevalence of missing BMI data through simple interventions such as sending measuring tapes or measuring charts to women in advance of the interview, although such methods would require validation.

In our study, missing data was an unlikely explanation for the positive association observed between prepregnancy obesity and spina bifida, although it could have impacted the magnitude of the observed association. The strong protective association between prepregnancy obesity and gastroschisis cannot be attributed to bias due to missing data. Our analysis did demonstrate, however, that when an observed association is null, such as for prepregnancy obesity and CLP, missing data could be obscuring an association.

In general, missing data could mask true non-null associations or create the appearance of associations when the true association is null; detailed knowledge of the missing data mechanism or extensive sensitivity analysis is needed on an analysis-by-analysis basis to provide information about the validity of any observed effects in a given analysis. Within the framework of Little and Rubin [16], a complete case analysis implicitly assumes that observations are missing completely at random (MCAR), whereby missingness is unrelated to any observed or unobserved factors, including outcome. Under an assumption of missing at random (MAR), missingness can be associated with factors, but only those about which information is available. In our analysis the multiple imputation and probabilistic simulation results assumed MAR. Our results assuming the data were MCAR and MAR were nearly identical, suggesting that missingness did not bias results under an assumption of MAR. The extreme scenarios we considered allowed us to consider the maximum possible impact of data missing not at random (MNAR).

Our study was subject to a few limitations. In NBDPS, height and weight were self-reported, which could result in either over- or under-estimating height and weight; however, given that we used BMI categories rather than height and weight directly, the results of our analysis were less likely to be affected. The option to systematically record height in centimeters was not available to interviewers until July 2002; however, it does not appear that this option was related to likelihood of missing BMI, overall or within strata of race/ethnicity (data not shown). For the probabilistic sensitivity analysis method, we did not incorporate case status into our probabilities; they were based solely on stratum-specific BMI category prevalence estimates from NHANES. These percentages are therefore based on the premise that case status is not related to BMI, which is the very association we are assessing, and is contradicted by our results for spina bifida and gastroschisis. Had we incorporated case status into the probabilities we would have been assuming the opposite, that case status is related to BMI; we therefore chose the more conservative option. Because of the relatively small fraction of missing data, we were able to examine all possible combinations of missing BMI categories; this would not be possible with larger missing data fractions or a variable that must be analyzed as continuous.

The strengths of our study include using a large population-based case-control study of birth defects with consistent and detailed case ascertainment and classification criteria. We utilized measured (rather than self-reported) height and weight data from an external data source to assign the probabilities for being in each BMI category based on factors associated with BMI. By estimating the “worst case” ORs we were able to put a bound on the possible values that could be observed if we had no missing data, while the multiple imputation and sensitivity analysis methods allowed us to estimate the most likely values. Although this paper presents data on BMI and specific birth defects, the different methods employed for assessing the potential impact of missing data demonstrate an application of the recommendations of a 2010 report by the National Research Council on the treatment of missing data in clinical trials, which also applies to observational studies (summarized in [31]).

Based on our findings we can conclude that the missing BMI in NBDPS should not be considered missing completely at random given that missingness can depend on observed characteristics, which introduces the potential for bias [16]. However, it does not appear that missing BMI data impacted conclusions about the presence or absence of an association from a traditional complete case analysis in these data, although it could impact the magnitude of the association. We have demonstrated simple methods for systematically and quantitatively estimating the potential impact of missing data which could be easily applied to other studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported through cooperative agreements under PA 96043, PA 02081 and FOA DD09-001 from the Centers for Disease Control and Prevention to the Centers for Birth Defects Research and Prevention participating in the National Birth Defects Prevention Study.

References

1. Adams KF, Schatzkin A, Harris TB, et al. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *The New England Journal of Medicine*. 2006; 355(8):763–778. [PubMed: 16926275]
2. Flegal KM, Carroll MD, Kit BK, et al. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999–2010. *JAMA*. 2012; 307(5):491–497. [PubMed: 22253363]
3. Flegal KM, Carroll MD, Kuczmarski RJ, et al. Overweight and obesity in the United States: Prevalence and trends, 1960–1994. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*. 1998; 22(1):39–47.
4. Flegal KM, Carroll MD, Ogden CL, et al. Prevalence and trends in obesity among US adults, 1999–2008. *JAMA*. 2010; 303(3):235–241. [PubMed: 20071471]
5. Must A, Spadano J, Coakley EH, et al. The disease burden associated with overweight and obesity. *JAMA*. 1999; 282(16):1523–1529. [PubMed: 10546691]
6. Ogden CL, Carroll MD, Kit BK, et al. Prevalence of obesity and trends in body mass index among US children and adolescents, 1999–2010. *JAMA*. 2012; 307(5):483–490. [PubMed: 22253364]
7. Ogden CL, Lamb MM, Carroll MD, et al. Obesity and socioeconomic status in adults: United States, 2005–2008. *NCHS Data Brief*. 2010; 50:1–8. [PubMed: 21211165]
8. Callaway LK, Prins JB, Chang AM, et al. The prevalence and impact of overweight and obesity in an Australian obstetric population. *The Medical Journal of Australia*. 2006; 184(2):56–59. [PubMed: 16411868]
9. Correa A, Marcinkevage J. Prepregnancy obesity and the risk of birth defects: An update. *Nutrition Reviews*. 2013; 71(Suppl 1):S68–S77. [PubMed: 24147927]
10. Sarwer DB, Allison KC, Gibbons LM, et al. Pregnancy and obesity: A review and agenda for future research. *Journal of Women's Health*. 2006; 15(6):720–733.
11. Stothard KJ, Tennant PW, Bell R, et al. Maternal overweight and obesity and the risk of congenital anomalies: A systematic review and meta-analysis. *JAMA*. 2009; 301(6):636–650. [PubMed: 19211471]
12. Waller DK, Shaw GM, Rasmussen SA, et al. Prepregnancy obesity as a risk factor for structural birth defects. *Archives of Pediatrics and Adolescent Medicine*. 2007; 161(8):745–750. [PubMed: 17679655]
13. Al-Lawati JA, Jousilahti P. Body mass index, waist circumference and waist-to-hip ratio cut-off points for categorisation of obesity among Omani Arabs. *Public Health Nutrition*. 2008; 11(1):102–108. [PubMed: 17572936]
14. Vazquez G, Duval S, Jacobs DR Jr, et al. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: A meta-analysis. *Epidemiologic Reviews*. 2007; 29:115–128. [PubMed: 17494056]
15. Barlow SE, Expert C. Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: Summary report. *Pediatrics*. 2007; 120(Suppl 4):S164–S192. [PubMed: 18055651]
16. Little, R.; Rubin, D. *Statistical analysis with missing data*. 2nd. New York: Wiley; 2002.
17. Engstrom JL, Paterson SA, Doherty A, et al. Accuracy of self-reported height and weight in women: An integrative review of the literature. *Journal of Midwifery & Women's Health*. 2003; 48(5):338–345.
18. Ramos E, Lopes C, Oliveira A, et al. Unawareness of weight and height—The effect on self-reported prevalence of overweight in a population-based study. *The Journal of Nutrition, Health and Aging*. 2009; 13(4):310–314.

19. Fonseca H, de Matos MG, Guerra A, et al. Emotional, behavioural and social correlates of missing values for BMI. *Archives of Disease in Childhood*. 2009; 94(2):104–109. [PubMed: 18801764]
20. Mikolajczyk RT, Richter M. Associations of behavioural, psychosocial and socioeconomic factors with over- and underweight among German adolescents. *International Journal of Public Health*. 2008; 53(4):214–220. [PubMed: 18716726]
21. Sherry B, Jefferds ME, Grummer-Strawn LM. Accuracy of adolescent self-report of height and weight in assessing overweight status: A literature review. *Archives of Pediatrics and Adolescent Medicine*. 2007; 161(12):1154–1161. [PubMed: 18056560]
22. Cogswell ME, Bitsko RH, Anderka M, et al. Control selection and participation in an ongoing, population-based, case-control study of birth defects: The National Birth Defects Prevention Study. *American Journal of Epidemiology*. 2009; 170(8):975–985. [PubMed: 19736223]
23. Yoon PW, Rasmussen SA, Lynberg MC, et al. The National Birth Defects Prevention Study. *Public Health Reports*. 2001; 116(Suppl 1):32–40. [PubMed: 11889273]
24. Rasmussen SA, Olney RS, Holmes LB, et al. Guidelines for case classification for the National Birth Defects Prevention Study. *Birth Defects Research, Part A: Clinical and Molecular Teratology*. 2003; 67(3):193–201.
25. National Institutes of Health. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults—The evidence report. National Institutes of Health. *Obesity Research*. 1998; 6(Suppl 2):51S–209S. [PubMed: 9813653]
26. World Health Organization. *Obesity: Preventing and managing the global epidemic*. Geneva: World Health Organization; 2000.
27. National Center for Health Statistics. Centers for Disease Control and Prevention National Health and Nutrition Examination Survey. 2012. (http://www.cdc.gov/nchs/nhanes/about_nhanes.htm). Accessed 3 Sept 2014
28. Centers for Disease Control and Prevention. National health and nutrition examination survey: Analytic guidelines, 1999–2010. *Vital and Health Statistics*. 2013; 2(161):1–24.
29. Rubin, D. *Multiple imputation for nonresponse in surveys*. New York: Wiley; 1987.
30. Molenberghs, G.; Kenward, MG. *Missing data in clinical studies*. New York: Wiley; 2007.
31. Little RJ, D’Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *The New England Journal of Medicine*. 2012; 367(14):1355–1360. [PubMed: 23034025]

Significance

There is no information on missing body mass index (BMI) and the factors associated with it in regards to maternal characteristics and birth defects. Furthermore, whether missing BMI has an effect on the observed association between obesity and certain birth defects is not understood. Our study aims to present multiple ways in handling missing BMI and how it would affect observed associations with certain birth defects.

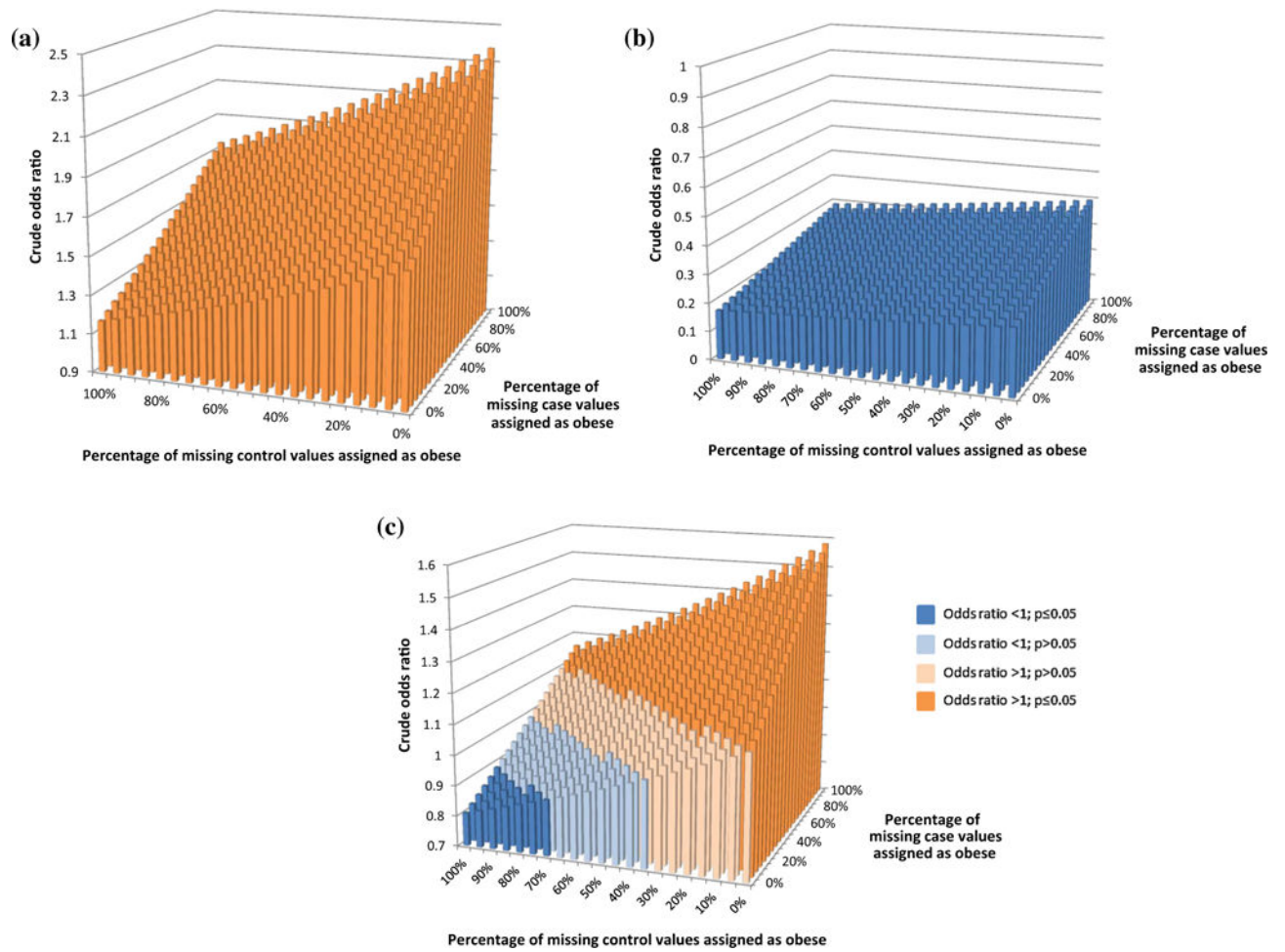


Fig. 1. Crude odds ratio for possible combinations of missing body mass index data for cases and controls for spina bifida, gastroschisis, and cleft lip with or without cleft palate. **a** Spina bifida. **b** Gastroschisis. **c** Cleft lip with or without cleft palate

Table 1

Characteristics associated with missing body mass index (calculated using height and weight) in the National Birth Defects Prevention Study control mothers (1997–2009)

	Missing N (%)	Not missing N (%)	Crude odds ratio	95 % CI	Adjusted odds ratio ^a	95 % CI
Total	435 (4.6)	9640 (95.4 %)				
Maternal age at EDD			Reference		Reference	
<18	18 (5.1)	333 (94.9)				0.8, 2.7
18–24	170 (5.8)	2821 (94.3)	1.1	0.7, 1.8	1.4	0.7, 2.6
25–29	125 (4.5)	2652 (95.5)	0.9	0.5, 1.4	1.4	0.7, 2.8
30–34	91 (3.6)	2468 (96.4)	0.7	0.4, 1.1	1.4	0.4, 1.8
35+	31 (2.2)	1366 (97.8)	0.4	0.2, 0.8	0.8	
Race/ethnicity						
Non-Hispanic white	32 (0.6)	5830 (99.4)	Reference		Reference	
Non-Hispanic black	7 (0.6)	1102 (99.4)	1.2	0.5, 2.6	0.3	0.1, 0.9
Hispanic	370 (15.6)	2010 (84.5)	33.5	23.3, 48.3	2.0	1.2, 3.4
Other	26 (3.6)	689 (96.4)	6.9	4.1, 11.6	1.1	0.6, 2.2
Maternal education						
<12 years	284 (17.0)	1384 (83.0)	5.7	4.4, 7.4	2.3	1.7, 3.1
12 years	82 (3.5)	2293 (96.5)	Reference		Reference	
Some college	24 (0.9)	2640 (99.1)	0.2	0.2, 0.4	0.5	0.3, 0.8
College graduate	15 (0.5)	3142 (99.5)	0.1	0.1, 0.2	0.4	0.2, 0.7
Language of interview						
English	152 (1.7)	8987 (98.3)	Reference		Reference	
Spanish	252 (29.2)	611 (70.8)	24.4	19.6, 30.3	2.4	1.8, 3.2
Country of birth						
USA	40 (0.5)	7790 (99.5)	Reference		Reference	
Other	365 (17.9)	1675 (82.1)	42.4	30.5, 59.1	8.6	5.5, 13.4
Parity						
First birth	124 (3.1)	3879 (96.9)	Reference		Reference	
Second birth	297 (4.9)	5758 (95.1)	1.6	1.3, 2.0	1.1	0.8, 1.5
Folic acid use (B1)						

	Missing N (%)	Not missing N (%)	Crude odds ratio	95 % CI	Adjusted odds ratio ^a	95 % CI
No	389 (5.7)	6465 (94.3)	Reference		Reference	
Yes	32 (1.0)	3134 (99.0)	0.2	0.1, 0.2	0.5	0.3, 0.7
Smoking (B1)						
No	392 (4.8)	7743 (95.2)	Reference		Reference	
Yes	18 (1.0)	1746 (99.0)	0.2	0.1, 0.3	0.7	0.4, 1.3
Alcohol (B1)						
No	372 (5.5)	6384 (94.5)	Reference		Reference	
Yes	36 (1.2)	3069 (98.8)	0.2	0.1, 0.3	0.7	0.4, 1.0

EDD estimated due date, B1 1 month before conception

^aEstimates are adjusted for all other variables presented in the table

Frequency of missing self-reported height, weight, and body mass index in the National Birth Defects Prevention Study control mothers (1997–2009) and non-pregnant women age 16–53 years in the National Health and Nutrition Examination Survey (1999–2010)

Table 2

	NBDPS control mothers 1997–2009 ^a			NHANES non-pregnant women 1999–2010 ^b				
	Total N	Missing height	Missing weight N (%)	Missing BMI	Total N	Missing height	Missing weight N (%) ^c	Missing BMI
All mothers	10,075	392 (3.9)	86 (0.8)	435 (4.6)	8420	198 (1.1)	183 (1.6)	319 (2.4)
Non-Hispanic white	5862	19 (0.3)	18 (0.3)	32 (0.5)	3469	4 (0.1)	51 (1.3)	53 (1.4)
Non-Hispanic black	1109	7 (0.6)	3 (0.3)	7 (0.6)	1966	11 (0.5)	31 (1.6)	35 (1.7)
Hispanic	2380	347 (14.6)	55 (2.3)	370 (15.6)	2029	152 (7.9)	86 (4.1)	190 (9.7)
Other race/ethnicity	715	19 (2.7)	10 (1.4)	26 (3.6)	956	31 (2.7)	15 (1.6)	41 (3.7)

BMI body mass index

^a Self-reported estimate of pre-pregnancy height and weight

^b Self-reported estimates of height and weight

^c Unweighted N and weighted percentage

Table 3

Potential impact of missing body mass index (BMI) on the association between BMI and selected birth defects utilizing four different methods

	I Complete case analysis				II Extreme scenarios				III Multiple imputation Median OR for Missing data	IV Probabilistic simulation Median OR for BMI category assigned to missing data
	Observed data		Missing deleted		Missing cases normal weight, missing controls obese		Missing cases obese, missing controls normal weight			
	Case	Control	Case	Control	Case	Control	Case	Control		
Spina bifida										
Normal weight (referent)	479	5229	479	5229	555	5229	479	5664		
Obese	261	1676	261	1676	261	2111	337	1676		
Missing	76	435	NA	NA	NA	NA	NA	NA	1.7	1.7
Crude odds ratio			1.7	1.7	1.2	2.4	2.4		1.5, 1.8	1.4, 2.0
95 % CI for crude odds ratios			1.5, 2.0	1.0, 1.4	1.0, 1.4	2.1, 2.8	2.1, 2.8		1.6	1.6
Adjusted ^a odds ratio			1.6	1.1	1.1	2.3	2.3		1.5, 1.8	1.3, 1.9
95 % CI for adjusted ^a odds ratios			1.4, 1.9	0.9, 1.3	0.9, 1.3	2.0, 2.7	2.0, 2.7			
Gastrochisis										
Normal weight (referent)	797	5229	797	5229	832	5229	797	5664		
Obese	58	1676	58	1676	58	2111	93	1676		
Missing	35	435	NA	NA	NA	NA	NA	NA	0.2	0.2
Crude odds ratio			0.2	0.2	0.2	0.4	0.4		0.0, 0.5	0.2, 0.3
95 % CI for crude odds ratios			0.2, 0.3	0.1, 0.2	0.1, 0.2	0.3, 0.5	0.3, 0.5		0.2	0.3
Adjusted ^a odds ratio			0.2	0.2	0.2	0.4	0.4		0.0, 0.5	0.2, 0.4
95 % CI for adjusted ^a odds ratios			0.2, 0.3	0.1, 0.2	0.1, 0.2	0.3, 0.5	0.3, 0.5			
Cleft lip with or without cleft palate										
Normal weight (referent)	1336	5148	1336	5148	1478	5148	1336	5581		
Obese	484	1653	484	1653	484	2086	626	1653		
Missing	142	433 ^b	NA	NA	NA	NA	NA	NA	1.1	1.1
Crude odds ratio			1.1	1.1	0.8	1.6	1.6		1.0, 1.2	1.0, 1.3
95 % CI for crude odds ratios			1.0, 1.3	0.7, 0.9	0.7, 0.9	1.4, 1.8	1.4, 1.8		1.1	1.1
Adjusted ^a odds ratio			1.1	0.8	0.8	1.6	1.6		1.0, 1.2	1.0, 1.3
95 % CI for adjusted ^a odds ratios			1.0, 1.3	0.7, 0.9	0.7, 0.9	1.4, 1.8	1.4, 1.8		1.0, 1.2	1.0, 1.3

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

OR odds ratio, *BMI* body mass index

^a Adjusted for maternal age, race/ethnicity, education, parity, smoking in the month prior to conception, and folic acid intake in the month prior to conception

^b The number of controls is different for cleft lip with or without cleft palate compared to the other defects because one of the NBDPS Centers did not submit cleft lip with or without cleft palate cases for 1 year of the study; therefore, controls from that Center for that year were not included in analysis of this defect