Throwing out the baby with the bathwater? Comparing two approaches to handling implausible values of change in body size

Christine L. Gray[1] and **Whitney R. Robinson**[1,2]

[1] University of North Carolina at Chapel Hill, Gillings School of Global Public Health, Department of Epidemiology

[2] University of North Carolina at Chapel Hill, Carolina Population Center

INTRODUCTION

The most commonly used measure of body size in epidemiologic studies of childhood obesity is body mass index (BMI). [1-5] The utility of BMI in childhood obesity research is dependent on accurate measurement of weight and height. One problem that can arise in analysis of childhood BMI is the appearance of height loss, or "shrinkage." In a population of growing children and adolescents, shrinkage is implausible and indicates that at least one height value is erroneous and should be corrected. However, ascertainment of the correct height values is often infeasible.

Sensitivity analysis and imputation are established methods for simulating correction of implausible data.[6-10] However, measurement error is rarely explicitly acknowledged and use of these methods is not yet widespread.[11] More commonly, shrinkage in a growing population is either 1) addressed by excluding shrinkers from analysis or 2) ignored, with shrinkers retained in analysis. Excluding shrinkers reduces sample size and therefore statistical precision, but precision loss may be a reasonable tradeoff if excluding shrinkers reduces bias from measurement error.[12] Ignoring the shrinkage has the advantage of retaining sample size but could result in biased estimates if retaining shrinkers introduces systematic measurement error.

To our knowledge, no studies using real-world epidemiologic data have evaluated whether excluding respondents with implausible height shrinkage results in more valid estimates than retaining shrinkers. Analyzing data on a nationally representative sample of older adolescents, we compared two strategies for handling data when some height-stable adolescents appear to shrink over a 6-year period between adolescence and young adulthood.

Corresponding Author: Christine L. Gray Department of Epidemiology, CB #7435 Gillings School of Global Public Health University of North Carolina at Chapel Hill Chapel Hill, NC 27599-7435 clgray@email.unc.edu Tel: 301.787.7634 Fax: 919.966.2089.

METHODS

Population

Data were from the National Longitudinal Study of Adolescent Health (Add Health) cohort. Add Health began as a nationally representative survey of US students enrolled in grades 7 through 12 in 1994–95.[13] Height and weight were self-reported at wave 1. At waves 2, 3, and 4, height and weight were both measured and self-reported.

To enable establishment of a gold standard height, we restricted the data to a subset of respondents with strong evidence of height stability across wave 2 (1996), wave 3 (2001-02) and wave 4 (2008-09). To identify height-stable respondents, we restricted to girls 17 years or older at wave 2. Approximately 95% of U.S. girls reach adult height by age 17, making the height-stability assumption more robust in adolescent females than males, who reach adult height later.[14] For height-stable female respondents, height should remain unchanged across all three waves. To establish a gold-standard height, we restricted to respondents whose measured height was the same at wave 3 and wave 4. If a respondent's wave 2 height was greater than the height consistently measured at waves 3 and 4, that respondent was considered a "shrinker." The final sample included 816 female respondents. All study procedures were approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Measures—To compare the two methods of handling implausible height shrinkage, we conducted illustrative analyses under three conditions: "gold-standard", excluding shrinkers, and retaining the shrinkers. The illustrative analyses were (a) estimation of wave 2 obesity prevalence and (b) estimation of the association between maternal education and wave 2 obesity prevalence. We dichotomized self-reported maternal education as ≤ high school and > high school for regression analyses.

Under all three conditions, wave 2 obesity was defined as either having BMI ≥ 30 kg/m^2 or being at or above the 95th percentile of the age-and sex-specific BMI specified by the CDC 2000 growth charts.[14-15] To construct a dataset with "gold standard" height, we substituted wave 2 measured height with the height value consistently measured at waves 3 and 4. Wave 3 height was measured in inches; wave 4 was measured in centimeters. Wave 4 height was considered consistent with wave 3 if it was within 0.5 inches of wave 3 height. eTable 1 shows variable distributions under each condition (see eTable 1, Supplemental Digital Content 1). Analyses were adjusted for Add Health's complex survey weighting using Stata 12 (College Station, Texas).[13,16]

Analysis—We compared estimated obesity prevalence when excluding or retaining shrinkers to the "gold standard" prevalence estimate. We also performed bivariate logistic regression analyses with maternal education as the independent variable under all three conditions ("gold standard," excluding shrinkers, retaining shrinkers).

To perform analyses under three conditions, we first calculated "gold-standard" estimates in the full analysis sample by substituting "naïve" wave 2 measured height with the consistent "gold standard" height. For the second and third conditions, we estimated obesity keeping

the "naïve" wave 2 measured height intact. In the second condition, shrinkers were excluded from the analysis sample. Shrinkers are those respondents whose consistent "gold standard" height was less than their "naïve" measured height at wave 2. In the third (naïve) condition, shrinkers were retained in the full analysis sample.

To compare estimates from the second and third conditions to the "gold standard," we created a stacked dataset that included the full "gold standard" analysis sample and duplicate rows for the non-shrinkers, i.e., the population after shrinkers were excluded. We used indicator variables to specify each analysis condition.

We produced bootstrapped estimates (2000 replications) corresponding to all analyses to assess the uncertainty of each reported estimate. Code to create the stacked dataset and conduct analyses is provided (see eAppendix: Analytic Code, Supplemental Digital Content 1).

RESULTS

As expected, excluding shrinkers resulted in a higher standard error (SE=1.9) than retaining shrinkers (SE=0.04) (Table 1). The difference was large, an order of magnitude. Notably, estimates from the sample that retained shrinkers were slightly closer to the gold standard estimate than those from the sample excluding shrinkers. "Gold standard" obesity prevalence was 14.0%. Obesity prevalence retaining and excluding shrinkers, respectively, was 13.8% and 13.4%.

Regression analyses also demonstrated worse precision when excluding shrinkers (see confidence limit ratios, Table 1). However, accuracy appeared much worse when excluding shrinkers versus retaining them. The odds ratio (OR) for the association of lower education and obesity was 2.21 using "gold standard" height (Table 1). When shrinkers were excluded, the OR dropped to 1.16. When shrinkers were retained, the OR (1.91) was closer to the "gold standard." An analysis excluding shrinkers would likely conclude there was no evidence that maternal education was associated with adolescent obesity, in contradiction to the literature in this area.[17] To further illustrate the differences between strategies, we produced histograms of bootstrapped regression coefficients under all three conditions (see Figure 1). Retaining shrinkers produced a distribution much closer to the "gold standard" distribution.

Additional analyses of mean BMI and bivariate associations with BMI showed that excluding and retaining shrinkers resulted in similar estimates close to the "gold standard." As expected, excluding shrinkers consistently produced lower precision. For all analyses, bootstrapped estimates supported our findings.

DISCUSSION

The primary rationale for excluding implausible values is to reduce bias. However, in our analyses, excluding respondents with implausible height change values did not reduce bias. Instead, excluding shrinkers simply reduced precision and even worsened accuracy in some instances. In contrast, ignoring the shrinkage and retaining shrinkers produced estimates

close to the "gold standard" estimates without loss of statistical precision. These findings suggest that removing shrinkers from the sample may, under some conditions, induce more bias than retaining those with implausible height change.

We believe that our results may be most relevant when measurement error is only partially observable. Further examination of our data indicated shrinkage was only one class of error in the data. Other error manifested as greater growth than actually occurred. However, in adolescent data, error is most manifest when shrinkage is observed. Excluding shrinkers removes observations from one extreme of the error distribution but retains observations affected by the remaining error. This selective exclusion may induce more bias than it prevents.

This study has several unique strengths. We designed and implemented an innovative study design. By exploiting the longitudinal design of Add Health, which followed adolescents into adulthood, we constructed a height-stable cohort and establish a "gold standard" height for older adolescents using their observed consistently measured adult height. Finally, we were able to examine a question of great practical interest in the important field of adolescent obesity research.

Our study is limited by restriction to a sample of older female adolescents. However, we have no reason to believe the pattern of measurement error differed by age or sex. We expect that the effect of handling shrinkage over time in a cohort of older adolescents would also apply in a younger population. Additionally, we assumed that consistently measured height at waves 3 and 4 represented "true" height. While true height is unknown, we believe consistently measured adult height represented an accurate measure for a sample of U.S. girls at least 17 years old.

In adolescent obesity research, the strategy of excluding shrinkers may not provide the validity gains researchers expect. Compared with retaining respondents who showed implausible shrinkage, excluding shrinkers did not improve accuracy and even diminished accuracy in some analyses. Excluding implausible height-change values may be intuitively appealing. However, our results suggest that ignoring the shrinkage and retaining shrinkers can be a less biased strategy than excluding them.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

REFERENCES

1. Dietz WH, Bellizzi MC. Introduction: the use of body mass index to assess obesity in children. Am J Clin Nutr. 1999; 70(1):123S–5S. [PubMed: 10419414]

2. Dietz WH, Robinson TN. Use of the body mass index (BMI) as a measure of overweight in children and adolescents. J Pediatr. 1998; 132(2):191–3. [PubMed: 9506622]

3. Barlow SE, Bobra SR, Elliott MB, et al. Recognition of childhood overweight during health supervision visits: Does BMI help pediatricians? Obesity. 2007; 15(1):225–32. [PubMed: 17228051]

4. Barlow SE, Dietz WH. Obesity evaluation and treatment: Expert Committee recommendations. The Maternal and Child Health Bureau, Health Resources and Services Administration and the Department of Health and Human Services. Pediatrics. 1998; 102(3):E29. [PubMed: 9724677]

5. Ogden CL, Yanovski SZ, Carroll MD, et al. The epidemiology of obesity. Gastroenterology. 2007; 132(6):2087–102. [PubMed: 17498505]

6. Greenland S. Basic methods for sensitivity analysis of biases. Int J Epidemiol. 1996; 25(6):1107–16. [PubMed: 9027513]

7. Lash TL, Fink AK. Semi-automated sensitivity analysis to assess systematic errors in observational data. Epidemiology. 2003; 14(4):451–8. [PubMed: 12843771]

8. Parker JD, Schenker N. Multiple imputation for national public-use datasets and its possible application for gestational age in United States Natality files. Paediatr Perinat Epidemiol. 2007; 21(Suppl 2):97–105. [PubMed: 17803623]

9. Klebanoff MA, Cole SR. Use of multiple imputation in the epidemiologic literature. Am J Epidemiol. 2008; 168(4):355–7. [PubMed: 18591202]

10. Spratt M, Carpenter J, Sterne JA, et al. Strategies for multiple imputation in longitudinal studies. Am J Epidemiol. 2010; 172(4):478–87. [PubMed: 20616200]

11. Jurek AM, Maldonado G, Greenland S, et al. Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. Eur J Epidemiol. 2006; 21(12):871–6. [PubMed: 17186399]

12. Rothman, KJ.; Greenland, S.; Lash, TL. Modern epidemiology. Third Edition. Wolters Kluwer Health/Lippincott Williams & Wilkins; Philadelphia: 2008.

13. Chantala, K. Internal Report. Carolina Population Center, University of North Carolina at Chapel Hill; 2006. Guidelines for analyzing Add Health data. http://www.cpc.unc.edu/projects/addhealth/data/guides/wt-guidelines.pdf. [June 6, 2013]

14. Ogden CL, Kuczmarski RJ, Flegal KM, et al. Centers for Disease Control and Prevention 2000 growth charts for the United States: improvements to the 1977 National Center for Health Statistics version. Pediatrics. 2002; 109(1):45–60. [PubMed: 11773541]

15. Kuczmarski RJ, Ogden CL, Guo SS, et al. CDC Growth Charts for the United States: methods and development. Vital Health Stat 11. 2000; 2002(246):1–190.

16. StataCorp. Stata Statistical Software: Release12. StataCorp LP; College Station, TX: 2011.

17. Classen T, Hokayem C. Childhood influences on youth obesity. Econ Hum Biol. 2005; 3(2):165–87. [PubMed: 15994141]
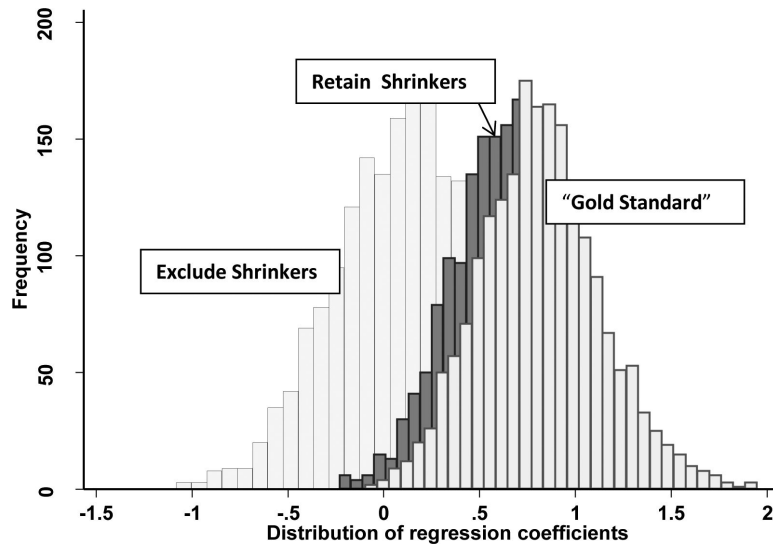
Figure 1.
Distribution of bootstrapped logistic regression coefficients under three conditions: "gold standard", excluding shrinkers, and retaining shrinkers.

Table 1

Estimates of obesity prevalence and bivariate association of obesity prevalence with maternal education, comparing two strategies of handling implausible height values to a "gold standard" estimate

	Obesity prevalence[a]			Association with maternal education	
	Prevalence estimate[b]	Difference from "Gold" estimate[b]	SE	OR (95% CI)[b]	Confidence Limit Ratio
"Gold Standard" Height[c] (N=816)	14.0	Ref	Ref	HS 2.21 (1.23, 3.97)	3.23
				>HS Ref	
Strategy 1: Exclude Shrinkers[d] (N=448)	13.4	−0.6	1.9	HS 1.16 (0.55, 2.46)	4.47
				>HS Ref	
Strategy 2: Retain Shrinkers[d] (N=816)	13.8	−0.2	0.4	HS 1.91 (1.07, 3.42)	3.20
				>HS Ref	

[a] Obesity is defined as BMI $\geq$ 30 kg/m^2 or being at or above the 95th percentile in weight-for-age according to the CDC 2000 growth chart. BMI is kg/m^2 using weight in pounds and height in inches.

[b] Prevalence and regression estimates account for the survey's complex sampling and use longitudinal sampling weights that account for loss-to-follow-up in waves 2, 3, and 4.

[c] Wave 2 height was substituted with the "gold standard" height that was consistent at waves 3 and 4 for analyses.

[d] Wave 2 height value is used for analyses.