



Published in final edited form as:

*Obesity (Silver Spring)*. 2016 May ; 24(5): 1132–1139. doi:10.1002/oby.21446.

## The prevalence and validity of high, biologically implausible values of weight, height and BMI among 8.8 million children

David S Freedman, PhD<sup>a</sup>, Hannah G Lawman, PhD<sup>b</sup>, Liping Pan, MD, MPH<sup>a</sup>, Asheley C Skinner, PhD<sup>c</sup>, David B Allison, PhD<sup>d</sup>, Lisa McGuire, PhD<sup>a</sup>, and Heidi M Blanck, PhD<sup>a</sup>

<sup>a</sup>Division of Nutrition, Physical Activity and Obesity, Centers for Disease Control and Prevention, Atlanta, GA

<sup>b</sup>National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD

<sup>c</sup>Department of Health Policy and Management, UNC School of Public Health, Chapel Hill, NC

<sup>d</sup>Nutrition Obesity Research Center, University of Alabama at Birmingham

### Abstract

**Objective**—We assessed the prevalence and consistency of high values of weight, height and BMI considered to be biologically implausible (BIV) using cut-points proposed by WHO among 8.8 million low-income children (13.7 million observations).

**Methods**—We performed cross-sectional and longitudinal analyses among 2- to 4-year-olds who were examined from 2008 through 2011.

**Results**—2.7% of the body size measurements were classified as biologically implausible; 95% of these BIVs were very high. Among the subset of children (3.6 million) examined more than once, most of those who initially had a high weight or BMI BIV also had a high BIV at the 2<sup>nd</sup> examination; odds ratios were >250. Based on several alternative classifications of BIVs, the current cut-points likely underestimate the prevalence of obesity by about 1%.

**Conclusions**—Many of the extremely high values of body size currently flagged as BIVs are unlikely to be errors. Increasing the z-score cut-points, or using a percentage of the maximum values in the National Health and Nutrition Examination Survey, could improve the balance between removing probable errors and retaining those that are likely correct.

### Introduction

The prevalence of obesity has increased by more than 50% among 2- to 5-year-olds since the early 1970s (1, 2). In response to this increase, a large amount of data on the weight and height of children have been collected in non-research settings, such as schools (3–8) and clinics (9–12). The quality of these measurements can vary substantially, and investigators

---

Address correspondence to: David S Freedman: CDC F-77, 4770 Buford Highway, Atlanta GA 30341 - 3724; phone: 770-488-6016, fax: 815-572-8152; dxfl@cdc.gov.

The views, findings and conclusions in this report are those of the authors and do not represent the official position of the Centers for Disease Control.

Disclosure: The authors declare no financial or other conflict of interest

must decide how to address the possible validity of extreme values. These decisions could influence prevalence estimates (8), secular trends, and comparisons across studies.

Weight, height, and BMI data that are considered to be biologically implausible (BIV) are flagged in the SAS program for the 2000 CDC growth charts (13, 14). These BIV cut-points were based on suggestions made by a WHO expert committee in the mid-1990s (15), and have been widely used (3, 4, 11, 12, 16–18). Values more extreme than the WHO cut-points were described as “likely to be errors and may be treated as missing” (p 217 of (15)). However, no justification was given for why these values were considered to be implausible, and the effects of these exclusions on prevalence estimates were not considered (15).

It is possible that the classification of values as biologically implausible may resemble other techniques that have been used to exclude ‘bad data’, but have eventually been found to introduce various biases (19–21). This possibility is supported by our recent analysis of data from Nutrition Examination Survey (NHANES) (22), which showed that most children who had a high BMI BIV also had arm and waist circumferences 99<sup>th</sup> percentile.

The objectives of the current study were (1) to evaluate the performance, based on repeated measurements of the same child, of the current cut points for high BIVs, and (2) to examine the effects of alternative BIV cut-points on the estimation of the prevalences of obesity and severe obesity. We use data from 2- to 4-year olds in the CDC’s Pediatric Nutrition Surveillance System (PedNSS) from 2008 through 2011 (N ~ 13.7 million records).

## Methods

### Study Sample

PedNSS was a state-based public health surveillance system that monitored the nutritional status from birth to age 5 y of low-income children who participated in federally funded maternal and child health and nutrition programs. PedNSS included almost 50% of children eligible for federally funded maternal and child health and nutrition programs, but more than 90% of children participated in the Special Supplemental Nutrition Program for Women, Infant, and Children (WIC) program.

Eligibility for WIC is based primarily (1) a current pregnancy or a child <5 years of age, and (2) a family income that is 185% of the federal income-poverty level (<http://www.census.gov/hhes/www/poverty/data/threshld/index.html>). Anthropometric measurements were taken by trained staff in state programs according to the CDC Nutrition Surveillance Program standards; weight was recorded to the nearest 1/4 pound and height to the nearest 1/8 inch with a measuring board (23). State agencies assigned a unique ID to each child.

There were 14.4 million records (observations) in the database from children who were between 24.0 and 59.9 months of age; many of these children were examined several times. The initial data cleaning excluded (1) 622,000 records that were missing data on weight or height, (2) 1357 records that had a value for weight or height of 0 or 999, and (3) 49,000 children for whom the birth date differed across records. These restrictions resulted in a total

of 13,690,110 records from 8,757,576 children. Children under 24 months of age were excluded from the analysis because both the index of relative weight and reference population (weight-for-length in the WHO growth standards) differs from those for older children. No other data cleaning was performed for the current analyses.

Based on the assigned ID, birthdate, and state, we were able to combine records for an individual children to create a subset (N=3,629,026) of children who were examined multiple times. (After this matching, it was found that the sex of 7367 (0.2%) children differed across examinations, and we excluded these children from the longitudinal analyses.) We focus on the first and second examinations for children who had multiple examinations, with the initial examination occurring at  $34 \pm 8$  months of age and the subsequent examination at  $45 \pm 8$  months. Because our focus is on extremely high values of weight, height and BMI, children who had a low BIV (N=17,737, 0.5%) at either examination were excluded from these analyses.

### Calculation of z-scores in the CDC growth charts

Body mass index (BMI,  $\text{kg}/\text{m}^2$ ) was calculated as weight divided by height<sup>2</sup>. Weight, height and BMI z-scores and percentiles were calculated using the CDC 2000 Growth Charts which are based on data collected from 1963–65 to 1988–94 (13). Obesity is defined as a BMI 95th percentile (13), and severe obesity as a BMI 120% of the 95th percentile (24).

Although most z-scores calculated in the CDC growth charts agree well with the empirical percentiles, estimates in the upper tails can be problematic (24–26). Therefore, a modified approach was used to identify extreme values in the SAS program for the CDC growth charts, with a child's 'modified z-score' based on  $\frac{1}{2}$  of the distance between 0 and 2 z-scores (25).

For example, a 49-month-old girl in the current analyses had a BMI of  $28.2 \text{ kg}/\text{m}^2$ . Based on the sex- and age-specific values of L (Box-Cox power transformation,  $-3.07$ ), M (mean, 15.28), and S (coefficient of variation, 0.079) in the CDC growth charts, her BMI-for-age z-score would be 3.5  $[(28.2 / M)^L - 1] / (L \cdot S)$  (27, 28). However, because  $\frac{1}{2}$  of the distance between 0 and 2 z-scores for this child is  $1.84 \text{ kg}/\text{m}^2$ , her modified z-score would be  $(28.2 - 15.28) / 1.84 = 7.0$ . (Because the BMI value at any z-score can be calculated as  $M \cdot [(1 + L \cdot S \cdot Z)^{(1/L)}]$ , a z-score of 2 would be  $18.96 \text{ kg}/\text{m}^2$ , and  $\frac{1}{2}$  of the distance between the BMI values at 0 and 2 z-scores is 1.84.) These modified z-scores extrapolate the SD distance based on non-extreme (absolute z-score of  $\leq 2$ ) values to more extreme values.

We refer to these modified z-scores as  $\text{weight}_{\text{MZ}}$ ,  $\text{height}_{\text{MZ}}$  and  $\text{BMI}_{\text{MZ}}$  or as a z-score<sub>M</sub>. The following extreme values are flagged as BIVs in the CDC SAS program (14):

$$\text{weight}_{\text{MZ}}: < -5 \text{ or } > +5$$

$$\text{height}_{\text{MZ}}: < -5 \text{ or } > +3$$

$$\text{BMI}_{\text{MZ}}: < -4 \text{ or } > +5$$

We consider body size measures below these cut-points to be a low BIV, while values above the cut points to be high BIVs. The focus in the current study is on high BIVs. These

modified z-scores are used only for BIVs, and not for the definitions of obesity and severe obesity.

### Statistical Analyses

Analyses were performed in R (29). We examined the z-score distributions using quantile-quantile (QQ) plots ([http://onlinestatbook.com/2/advanced\\_graphs/q-q\\_plots.html](http://onlinestatbook.com/2/advanced_graphs/q-q_plots.html)), in which the ordered quantiles (observed z-scores) were plotted vs. the expected quantiles from a normal distribution.

In the longitudinal cohort, we assessed the cross-classification high BIVs at baseline and follow-up. If high BIVs were data entry errors, one would expect that the probability of having a high BIV at the second examination would not be influenced by the child's values at the first examination. Associations between BIVs at the initial and subsequent examinations were summarized using odds ratios (OR) and positive likelihood ratios.

To assess the effects of BIVs on the estimated prevalence of obesity and severe obesity, we used several alternative cut-points for extremely high values, including weight and BMI z-score<sub>MS</sub> of +8 and +15 (and height z-score<sub>MS</sub> of +4 or +5). These cut points were based on an examination of the proportion of children in each 1-unit z-score category that had a high BIV at the second examination and on our previous analyses of NHANES 1999–2000 through 2011–12 data (22).

### Results

Descriptive characteristics of the children are shown in Table 1. (The 13.7 million observations in Table 1 are from 8.8 million children). Overall, 40% were Hispanic, 33% were non-Hispanic (NH) white, 19% were NH black, and 4% were classified in other categories or were missing information. The prevalence of obesity ranged from 12% (NH whites and blacks) to 20% (American Indians), while the prevalence of severe obesity varied from 0.9% to 1.6%. Overall, 2.7% of examination records had a body size measure that was considered to be biologically implausible, and 95% of these records had a high BIV. Of the 357,000 high BIV records, 28% had a high weight, 48% had a high height and 52% had a high BMI. Of the 185,000 children who had a high BMI<sub>MZ</sub>, 43% also had a high weight<sub>MZ</sub> and 6% had a high height<sub>MZ</sub>. Although prevalences of high BIVs for weight and BMI were highest among Hispanics and American Indians, and were higher among boys than girls, black children had the highest prevalence of high height BIVs.

Figure 1 shows the distribution of the modified z-scores for weight (top panel), height and BMI (right panel) for a random sample of 100,000 observations. Whereas a QQ plot of a normal distribution would be a diagonal line, the body size measures had very heavy tails. As assessed by estimates of skewness and kurtosis for each variable, the distribution of height was closest to that of a normal distribution, while BMI had the heaviest tails; 0.25% of the children had a BMI<sub>MZ</sub> > 50. Based on an inspection of the weight and height values of these very high z-scores, in many cases it appeared that (1) a decimal had been misplaced or (2) weight and height values had been interchanged.

We then focused on the 3.6 million children who were examined twice between the ages of 24 and 59 months. Table 2 shows the percentages of children in various baseline categories of weight<sub>MZ</sub>, height<sub>MZ</sub> or BMI<sub>MZ</sub> who had a high BIV at follow-up; for each measure, the upper range of plausible values based on the current cut-points is bolded. For example, of the 1.4 million children who had a baseline weight<sub>MZ</sub> below 0 (left column, 1<sup>st</sup> row), only 0.1% had a follow-up weight<sub>MZ</sub> >5. In contrast, more than 50% of children in each baseline weight<sub>MZ</sub> category between 6.0 and 14.9 had a high weight BIV at follow-up. At higher levels of weight<sub>MZ</sub>, however, the prevalence of high, follow-up BIVs decreased, suggesting that many of the weight<sub>MZ</sub> values >15 were errors.

The prevalences of a high height BIV (middle columns) at follow-up were lower than those for a high weight<sub>MZ</sub>, but children with an initial height<sub>MZ</sub> of 4.0 to 5.9 were about 5 times more likely to subsequently have a high BIV than were children with an initial z-score of 2 to 2.9. The right columns of Table 2 focus on BMI<sub>MZ</sub> levels. Whereas 30% of the 28,000 children who had an initial BMI<sub>MZ</sub> of 4.0 to 4.9 had a high BMI BIV at follow-up, prevalences were >60% for initial values between 6.0 and 14.9. The prevalence of a high BMI BIV at follow-up, however, was only 8% among children who had initial BMI<sub>MZ</sub> >25.

Table 3 shows levels of BMI, weight and height for 20 randomly selected children in longitudinal subset who had an initial BMI<sub>MZ</sub> between 12.5 and 14.9. (This range was selected because it was the highest BMI<sub>MZ</sub> category in Table 2 in which the positive predictive value was >50%). The first 12 children had a BMI<sub>MZ</sub> > 10 at follow-up, and 5 children had a follow-up z-score that was larger than the initial z-score. The last 5 children had follow-up values of BMI<sub>MZ</sub> that were fairly close to 0 suggesting that either the initial or follow-up measurement was an error (or that records were incorrectly matched).

Table 4 shows additional information on the cross-classification of initial and follow-up BIVs. The cross-classification of high BMI BIVs at the 2 examinations (bottom rows), for example, resulted in an OR of 254, a positive predictive value of 64%, and a positive likelihood ratio of 132. In general, the specificities, ORs and positive likelihood ratios were very high, but the sensitivities were only moderate (21% to 43%). Overall, the ORs for the weight and BMI associations were about 10-fold larger than that for height. Although the observed associations varied somewhat by sex, these differences were small relatively to the magnitudes of the associations (data not shown).

Table 5 shows the prevalences of obesity and severe obesity by race-ethnicity and sex based on the current (first set of columns) and several alternative BIV classifications. As seen in the 2<sup>nd</sup> set of columns, increasing the upper z-score cut-points to +8 (weight and BMI) and +4 (height) increased the prevalence of both obesity and severe obesity by 0.9% based on the inclusion of 228,000 additional records. In contrast, if the upper cut points were either (1) increased to z-scores of +15 (weight and BMI) and + 5 height (3<sup>rd</sup> set of columns) or (2) based on the age-specific maximum values observed in NHANES (1999–2000 through 2011–12, slightly more children were added the analysis, but the prevalence estimates varied by only 0.1 to 0.2 between these 3 classifications.

Analyses stratified by sex and race-ethnicity indicated that, in general, the prevalence increases with higher cut-points were larger among boys (+1.1% for both obesity and severe obesity for z-scores of +8 and +4) than among girls (about 0.7%), and that the increases were largest among Hispanic and American Indian / Alaskan Native boys. Among Hispanic boys, increasing the z-score cut-points to +8 (BMI and weight) and +4 (height) increased the prevalence of severe obesity from 1.3% to 2.9%. Furthermore, the prevalence of severe obesity was higher among girls than among boys (1.4% vs. 0.9%) based on the current BIV cut-points, but the prevalences were almost identical among boys and girls using any of the expanded cut-points. This reflects the higher prevalence of high BMI BIVs (based on the current cut-points) among boys than among girls (1.6% vs. 1.1%).

## Discussion

Our results indicate that among 2- to 4-year-olds in PedNSS, many of the high values of weight, height and BMI considered to be biologically implausible are unlikely to be errors. Of the 13.7 million observations in the current analyses, 2.6% had a high BIV based on the 1995 WHO cut-points (14, 15). Based on the longitudinal relation of levels of these body size measures among 3.6 million children, there were very strong associations between high BIVs at 2 examinations for BMI, weight and height. As compared with the estimated prevalences based on the current WHO cut-points, increasing the BIV cut-points, either through the use of higher z-scores or by using the age-specific NHANES maximums for weight, height and BMI, increased the prevalence of both obesity and severe obesity by about 1%.

The classification of BIVs in anthropometry was addressed in a 1995 WHO report (15) which recommended that z-scores  $> 5$  (weight and weight-for-height) and  $> 3$  (height) be treated as improbable and coded as missing. Although little justification was given for these specific values, it was stated that they were the “exclusion criteria for anthropometric values that are most likely to represent errors” (15) (p 218). These values may have been selected, in part, based on the low probability of values this extreme in a normal  $N(0,1)$  distribution. However, in both the current study and in NHANES (22), the BMI and weight distributions had very heavy right tails. Among 2- to 5-year-olds in NHANES (22), for example, the prevalence of a high BMI BIV (0.4%) was about 12,000 times than that of a normally distributed variable. The higher prevalence (2.6%) in the current study may, in part, be due to the higher prevalence of obesity among children in PedNSS and a larger number of coding errors. In general, the race-ethnicity and sex differences in the prevalence of high BMI BIVs paralleled differences in the prevalence of obesity.

The CDC SAS program to calculate BMI z-scores and percentiles (14), which has been widely used (3, 4, 11, 12, 16, 17), flags values considered to be implausible based on the 1995 WHO cut-points. Previous investigators have noted that these cut-points can result in a large number of extremely high values (16, 30) being classified as BIVs even though many are unlikely to be errors. We have previously shown (22) that most children in NHANES with a high BMI BIV had arm and waist circumferences that were  $> 99^{\text{th}}$  percentile. The longitudinal analyses in the current study provide further evidence that many of these high



BIVs are unlikely to be data errors. The concept of implausible values in other areas has also been questioned (31).

In addition to expanding the range of plausible values, based on z-scores or NHANES maximum values as in the current study, other approaches, such as the use of studentized residuals (22), may be useful. For example, rather than assessing the implausibility of BMIs without considering other data, an examination of regression residuals would allow one to assess whether a BMI is unusual given the child's sex, age, waist circumference, or other values (22, 32). In addition, some investigators (33–35) have suggested that values that are considered to be unlikely set to missing and then estimated using multiple imputation techniques (36). These methods, however, would require additional measurements that may not be available.

Although we did not focus on low BIVs, it is likely that many of these very low values were errors. The prevalence of low BIVs in the current study was 0.7% whereas only 2 of 5919 2- to 5-year-olds in NHANES 1999–2012 had a low BIV (22). This may be, in part, due to the relatively high number of high height BIVs that result in very low BMIs. In addition, longitudinal analyses in the current study also indicated (data not shown) that only 7% of the children who had a low weight BIV at the initial examination had a low weight BIV at a subsequent examination, suggesting the most of these low BIVs were errors. Of the ~90,000 records in the current study that had a low BIV, 28,000 had a height < 40 cm and 6200 had a weight < 5 kg. These values are less than 50% of the age-specific, minimums in NHANES 1999–2012 and are almost certainly errors.

Although the current cut-points for BIVs may incorrectly flag many of the extremely high values as implausible, the optimal solution may depend upon the analysis and the dataset. For data that have been very carefully recorded and cleaned, such as NHANES, it is likely that very few, if any, BIVs are errors. Although there is a tradeoff between including extremely high values that are correct and excluding those that are errors, our results suggest that even in data that has not been collected and managed as carefully, many of the values that are currently considered to be 'implausible' are unlikely to be errors. However, in the estimation of obesity and severe obesity prevalences, the choice of the upper cut point for an expanded range makes relatively little difference (Table 5) because most values that are currently considered to be implausible tend to cluster near the upper WHO cut-points. Changing the z-score cut points from +8 to +15, for example, results in relatively few additional children being included in a calculation.

Several limitations of our analyses should be considered. Although it is likely that many of the high BIVs that we observed were likely to be correct, this may not be true in all datasets. It should also be realized that data errors can occur throughout the entire distribution (32) and our analyses do not address this aspect of data cleaning. (These 'inliers' could be detected by an examination of studentized residuals.) It is possible, however, that the magnitude of the longitudinal association between extremely high values of the body size measures across the 2 examinations is even stronger than what we observed due to (1) mistakes in the matching process used to construct the cohort and (2) regression to the mean (37). The latter would result in follow-up values that would be expected to be less extreme

than the initial value. Although we think it unlikely, it is also possible that a high BIV at both examinations could represent either a measurement or transcription error at both examinations.

Our results indicate that simply deleting data based on the current BIV guidelines excludes a large number of high values that are likely to be correct. Although the current BIV cut-points could be substantially expanded, other data from the child, such as circumferences, could also be used to assess the validity of high BIVs. As compared with the WHO cut-points, the use of higher BIV cut-points would provide a better balance between excluding extremely high values that are likely to be errors and retaining those that are likely to be correct.

## Acknowledgments

We thank Ray Smith for his contribution to PedNSS data management and Patricia Brindley for her contribution to nutrition surveillance system operability.

## Abbreviations

<b>BIV</b>	biologically implausible value
<b>BMI</b>	body mass index
<b>CDC</b>	Centers for Disease Control and Prevention
<b>NHANES</b>	National Health and Nutrition Examination Survey
<b>PedNSS</b>	Pediatric Surveillance Nutrition Surveillance System
<b>WIC</b>	Women, Infants and Children Word

## References

1. Ogden CL, Flegal K, Carroll MD, Johnson CL. Prevalence and trends in overweight among US children and adolescents, 1999–2000. *JAMA*. 2002; 288:1728–1732. [PubMed: 12365956]
2. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA*. 2014; 311:806–14. [PubMed: 24570244]
3. Berger M. Obesity in K-8 students - New York City, 2006–07 to 2010–11 school years. *MMWR Morb Mortal Wkly Rep*. 2011; 60:1673–8. [PubMed: 22169977]
4. Kim J, Must A, Fitzmaurice GM, et al. Incidence and remission rates of overweight among children aged 5 to 13 years in a district-wide school surveillance system. *Am J Public Health*. 2005; 95:1588–94. [PubMed: 16051932]
5. Justus MB, Ryan KW, Rockenbach J, Katterapalli C, Card-Higginson P. Lessons learned while implementing a legislated school policy: body mass index assessments among Arkansas's public school students. *J Sch Health*. 2007; 77:706–713. [PubMed: 18076417]
6. Weedn AE, Ang SC, Zeman CL, Darden PM. Obesity prevalence in low-income preschool children in Oklahoma. *Clin Pediatr (Phila)*. 2012; 51:917–22. [PubMed: 22523274]
7. Robbins JM, Mallya G, Polansky M, Schwarz DF. Prevalence, Disparities, and Trends in Obesity and Severe Obesity Among Students in the Philadelphia, Pennsylvania, School District, 2006–2010. *Prev Chronic Dis*. 2012; 9:1–11.
8. Day SE, Konty KJ, Leventer-Roberts M, Nonas C, Harris TG. Severe obesity among children in New York City public elementary and middle schools, school years 2006–07 through 2010–11. *Prev Chronic Dis*. 2014; 11:E118. [PubMed: 25011000]



9. Pan L, May AL, Wethington H, Dalenius K, Grummer-Strawn LM. Incidence of obesity among young U.S. children living in low-income families, 2008–2011. *Pediatrics*. 2013; 132:1006–13. [PubMed: 24276843]
10. Smith N, Coleman KJ, Lawrence JM, et al. Body weight and height data in electronic medical records of children. *Int J Pediatr Obes*. 2010; 5:237–42. [PubMed: 19961272]
11. Novotny R, Oshiro CES, Wilkens LR. Prevalence of Childhood Obesity among Young Multiethnic Children from a Health Maintenance Organization in Hawaii. *Child Obes*. 2013; 9:35–42. [PubMed: 23373877]
12. Sekhobo JP, Edmunds LS, Reynolds DK, Dalenius K, Sharma A. Trends in prevalence of obesity and overweight among children enrolled in the New York State WIC program, 2002–2007. *Public Health Rep*. 2010; 125:218–24. [PubMed: 20297748]
13. Kuczumski RJ, Ogden CL, Guo SS, et al. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat 11*. 2002; 11:1–190. [PubMed: 12043359]
14. Centers for Disease Control and Prevention (CDC)., Centers for Disease Control and Prevention. A SAS program for the 2000 CDC growth charts (ages 0 to <20 years). 2015. [WWW document]. URL <http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm>
15. WHO Expert Committee. Physical status: the use and interpretation of anthropometry. WHO Tech Rep Ser. 1995:217–250.
16. Dennison BA, Edmunds LS, Stratton HH, Pruzek RM. Rapid infant weight gain predicts childhood overweight. *Obesity (Silver Spring)*. 2006; 14:491–9. [PubMed: 16648621]
17. Gundersen C, Lohman BJ, Eisenmann JC, Garasky S, Stewart SD. Child-specific food insecurity and overweight are not associated in a sample of 10- to 15-year-old low-income youth. *J Nutr*. 2008; 138:371–8. [PubMed: 18203906]
18. Kann L, Kinchen S, Shanklin SL, et al. Youth risk behavior surveillance--United States, 2013. *MMWR Surveill Summ*. 2014; 63(Suppl 4):1–168.
19. Allison DB, Heo M, Flanders DW, Faith MS, Williamson DF. Examination of “early mortality exclusion” as an approach to control for confounding by occult disease in epidemiologic studies of mortality risk factors. *Am J Epidemiol*. 1997; 146:672–680. [PubMed: 9345122]
20. Rhee JJ, Sampson L, Cho E, Hughes MD, Hu FB, Willett WC. Comparison of methods to account for implausible reporting of energy intake in epidemiologic studies. *Am J Epidemiol*. 2015; 181:225–33. [PubMed: 25656533]
21. Gray CL, Robinson WR. Throwing out the baby with the bathwater?: Comparing 2 approaches to implausible values of change in body size. *Epidemiology*. 2014; 25:591–4. [PubMed: 24809955]
22. Freedman DS, Lawman HG, Skinner AC, McGuire LC, Allison DB, Ogden CL. Validity of the WHO cutoffs for biologically implausible values of weight, height, and BMI in children and adolescents in NHANES from 1999 through 2012. *Am J Clin Nutr*. 2015; 102:1000–6. [PubMed: 26377160]
23. Mei Z, Scanlon KS, Grummer-Strawn LM, Freedman DS, Yip R, Trowbridge FL. Increasing prevalence of overweight among US low-income preschool children: the Centers for Disease Control and Prevention pediatric nutrition surveillance, 1983 to 1995. *Pediatrics*. 1998; 101:E12. [PubMed: 9417176]
24. Flegal KM, Wei R, Ogden CL, Freedman DS, Johnson CL, Curtin LR. Characterizing extreme values of body mass index-for-age by using the 2000 Centers for Disease Control and Prevention growth charts. *Am J Clin Nutr*. 2009; 90:1314–1320. [PubMed: 19776142]
25. Centers for Disease Control and Prevention. Cut-offs to define outliers in the 2000 CDC Growth Charts. [WWW document]. URL <http://www.cdc.gov/nccdphp/dnpa/growthcharts/resources/BIV-cutoffs.pdf>
26. Flegal KM, Cole TJ. Construction of LMS parameters for the Centers for Disease Control and Prevention 2000 growth charts. *Natl Health Stat Report*. 2013; 9:1–3. [PubMed: 24992748]
27. Cole T, Bellizzi M, Flegal K, Dietz W. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*. 2000; 320:1240–3. [PubMed: 10797032]
28. Cole TJ. The development of growth references and growth charts. *Ann Hum Biol*. 2012; 39:382–94. [PubMed: 22780429]

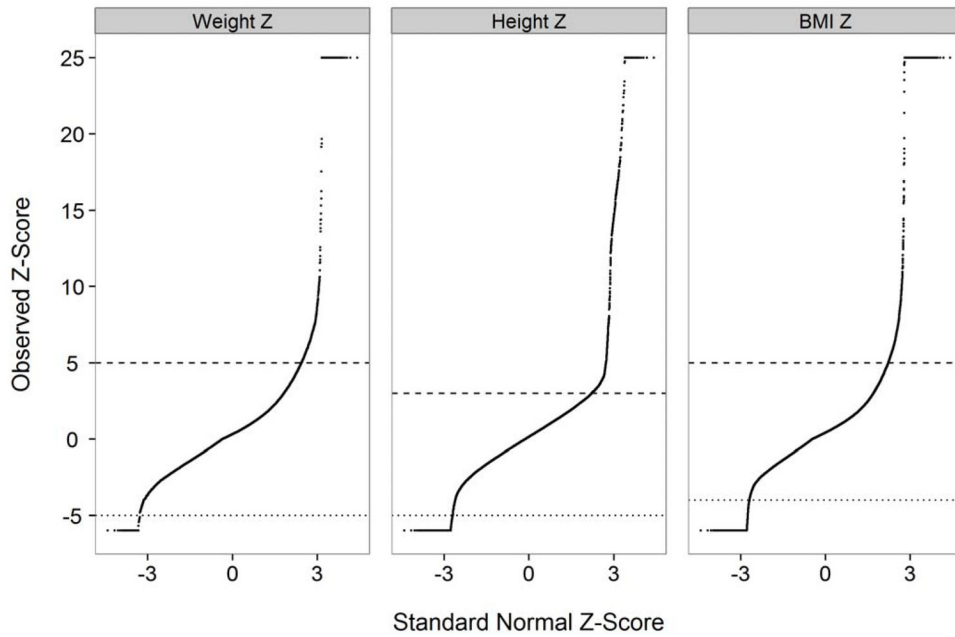
29. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: [WWW document]. URL <http://www.r-project.org/>
30. Lloyd K, Ohri-Vachaspati P, Chou J. Biologically implausible values and childhood obesity surveillance. *FASEB J*. 2011; 25:215.1.
31. Gaudino JA, Blackmore-Prince C, Yip R, Rochat RW. Quality assessment of fetal death records in Georgia: a method for improvement. *Am J Public Health*. 1997; 87:1323–7. [PubMed: 9279268]
32. Van den Broeck J, Argeseanu Cunningham S, Eeckels R, Herbst K. Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities. *PLoS Med*. 2005; 2:e267. [PubMed: 16138788]
33. 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]
34. Little, RJA.; Smith, PJ. *Editing and Imputation for Quantitative Survey Data*. 2015. [WWW document]. URL [http://www.jstor.org/stable/2289125?seq=1#page\\_scan\\_tab\\_contents](http://www.jstor.org/stable/2289125?seq=1#page_scan_tab_contents)
35. Hediger ML, Overpeck MD, McGlynn A, Kuczmarski RJ, Maurer KR, Davis WW. Growth and fatness at three to six years of age of children born small- or large-for-gestational age. *Pediatrics*. 1999; 104:e33. [PubMed: 10469816]
36. Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM, Donders TAR. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol*. 2006; 59:1087–1091. [PubMed: 16980149]
37. Bland JM, Altman DG. Regression towards the mean. *BMJ*. 1994; 308:1499. [PubMed: 8019287]

**What is known about this subject?**

1. The WHO cut-points for biologically implausible values are widely used in conjunction with the CDC growth charts for levels of weight, height, and BMI among children and adolescents
2. There is some evidence that these cut-points may exclude a large number of children who have very high, but correct, values of the body size measures

**What this study adds**

1. We estimate the prevalence of high BIVs among 8.8 million 2- to 4-year-olds
2. Based on the results of longitudinal analysis in a subset of 3.6 million, we show that most of the children with BIVs for weight and BMI at their initial examination also have values that would be classified as BIVs at the 2<sup>nd</sup> examination.
3. Our results suggest that substantially expanding the BIV upper cut-points could improve the balance between removing probable errors and retaining those that are likely correct.



**Figure 1.**

QQ plots of modified z-scores for weight (left), height, and BMI (right) among 100,000 randomly selected children in PedNSS. The horizontal lines represent the cut-points for values currently considered to be a high BIV (+ 5 for weight and BMI; +3 for height) or a low BIV (−5 for weight and height; −4 for BMI). Modified z-scores below −6 were recoded to −6 and values >25 were recoded to 25 to focus on the regions of interest. If the data were normally distributed, all points would fall on a diagonal line. In contrast, all body size measures had heavy tails at both ends of the distribution. Even within the range of values currently considered to be plausible, the distributions of weight and BMI were more skewed and kurtotic than a normal distribution.

**Table 1**

Descriptive characteristics of children in the PedNSS (2008 to 2011), by race-ethnicity.

Characteristic	Overall	Race-Ethnicity <sup>d</sup>					
		White	Black	Hispanic	Asian / Pacific Islander	American Indian / Alaskan Native	
N	13,690,110	4,526,673	2,666,678	5,426,963	360,743	105,743	
Boys (%)	50.66 <sup>b</sup>	51.0	50.6	50.4	50.8	50.5	
Girls (%)	49.34	49.0	49.4	49.6	49.2	49.5	
Age 2 y (%)	37.24	38.0	38.1	36.0	36.6	36.8	
Age 3 y (%)	32.59	32.8	32.8	32.5	32.1	32.7	
Age 4 y (%)	30.17	29.2	29.1	31.5	31.4	30.5	
BIV (%) <sup>c</sup>	2.75	1.84	2.75	3.30	2.45	3.14	
High BIV (%) <sup>c</sup>	2.61	1.72	2.63	3.15	2.30	3.01	
Low BIV (%) <sup>c</sup>	0.67	0.34	0.35	0.92	0.75	0.32	
High BMI BIV (%)	1.35	0.92	0.92	1.83	1.06	1.64	
BMI ≥ 85 (%) <sup>d</sup>	30.71	28.2	26.2	35.3	25.5	39.8	
Obesity (%) <sup>d</sup>	14.44	12.3	11.7	17.8	11.6	20.0	
Severe Obesity (%) <sup>d</sup>	1.18	0.85	0.86	1.64	0.89	1.59	

<sup>a</sup>Race-specific estimates exclude observations whose race was missing, as well as those children who were classified as another category or as 'unknown' race. Values represent observations, not children

<sup>b</sup>Values are column percentages.

<sup>c</sup>Current z-score values that are considered to be BIVs are modified BMI z-scores < -4 or > 5, modified weight z-scores < -5 or > 5, and modified height z-scores < -5 or > 3.

<sup>d</sup>Limited to participants without a BIV. Severe obesity is defined as a BMI-for-age 120% of the 95<sup>th</sup> percentile

**Table 2**

Relation of modified z-scores at initial examination to high BIVs for weight, height and BMI at the 2<sup>nd</sup> examination among 3,606,099 children

Modified z-score at Initial Examination	Weight <sup>d</sup>		Height <sup>d</sup>		BMI <sup>d</sup>	
	N (Initial Exam)	High BIV at 2 <sup>nd</sup> Exam (%)	N (Initial Exam)	High BIV at 2 <sup>nd</sup> Exam (%)	N (Initial Exam)	High BIV at 2 <sup>nd</sup> Exam (%)
<0	1,411,342 <sup>b</sup>	0.1 <sup>b</sup>	1,754,219	0.2	1,196,831	0.1
0 to 1.9	1,892,841	0.2	1,645,768	1.0	2,082,434	0.3
2.0 to 2.9	191,767	1.6	<b>165,768</b>	<b>6.8</b>	203,227	2.5
3.0 to 3.9	63,267	7.0	33,555	17.9	67,291	11.0
4.0 to 4.9	<b>24,608</b>	<b>20.8</b>	5,270	32.8	<b>27,885</b>	<b>29.7</b>
5.0 to 5.9	10,711	43.7	910	35.4	12,994	54.7
6.0 to 6.9	4,712	67.9	264	17.0	6,314	71.1
7.0 to 7.9	2,145	83.5	115	12.2	3,218	76.2
8.0 to 9.9	1,613	89.2	99	17.2	2,427	80.6
10 to 12.4	440	84.5	30	40.0	724	80.8
12.5 to 14.9	113	58.4	29	6.9	178	60.1
15.0 to 17.4	53	18.9	22	0.0	84	39.3
17.5 to 19.9	35	8.6	17	0.0	51	23.5
20 to 24.9	39	2.6	14	7.1	30	13.3
25	2,413	7.7	19	10.5	2,411	8.3

<sup>a</sup>Current guidelines consider the upper cut-points of valid body size measures to be modified z-scores < 5 (weight and BMI) and < 3 (height). Values above these ranges are considered to be high BIVs. For each body size measure, the upper range of plausible values based on the current cut-points is bolded. Children with a low BIV have been excluded from all analyses

<sup>b</sup>Values represent Ns at the 1<sup>st</sup> examination and percent of these children who had a high BIV at the 2<sup>nd</sup> examination



**Table 3**  
 Body size measures of 20 (of 178) children who had an initial modified BMI z-score of 12.5 to 14.9<sup>a</sup>

	Modified BMI z-score		BMI		Weight		Height		Age (months)	
	Initial	Follow-up	Initial	Follow-up	Initial	Follow-up	Initial	Follow-up	Initial	Follow-up
1	13.4	16.7	35.3	38.8	32.9	38.1	97	99	35	41
2	14.7	15.8	35.9	38.4	42.8	49.6	109	114	45	51
3	14.2	14.7	35.4	36.1	44.2	47.1	112	114	41	47
4	14.1	14.6	35.1	37.2	42.9	49.4	111	115	47	54
5	13.4	14.4	35.2	35.8	39.1	45.9	105	113	36	48
6	13.8	13.8	34.8	34.8	45.4	45.4	114	114	41	44
7	14.8	13.4	36.1	34.8	44.5	46.2	111	115	45	51
8	13.4	13.4	34.2	34.4	33.6	41.5	99	110	41	49
9	13.2	12.9	33.8	33.8	43.4	43.4	113	113	44	49
10	13.5	12.4	34.3	33.1	35.8	40.9	102	111	43	49
11	12.8	11.7	33.3	33.4	43.5	51.7	114	125	44	56
12	13.3	10.8	34.4	30.9	37.3	40.4	104	114	38	49
13	13.0	6.3	33.6	24.5	28.1	32.0	91	114	41	50
14	14.6	5.8	36.9	24.7	25.9	28.1	84	107	52	58
15	14.0	4.3	38.9	22.8	23.7	26.0	78	107	36	42
16	13.7	2.2	39.2	19.7	44.6	26.3	107	116	44	59
17	14.4	1.9	39.7	18.8	<b>172.4</b>	19.4	<b>208</b>	102	40	51
18	13.6	1.3	38.1	17.6	41.3	24.0	104	117	37	48
19	14.9	-0.9	37.0	14.7	37.2	17.2	100	108	37	49
20	15.0	-1.2	41.5	14.5	36.1	14.6	93	100	25	37

<sup>a</sup>Children are arranged in descending order of modified BMI z-score at follow-up. Values that are clearly errors are bolded.

**Table 4**

Cross-classification of high BIVs<sup>a</sup> at the initial and subsequent examination

High, Initial BIV	High BIV at age 4 y				Classification Performance				
	No	Yes	Row Sum	Odds Ratio (95% CI) <sup>b</sup>	+ Predictive value	Sensitivity	Specificity	+ Likelihood Ratio	
Modified weight									
z-score of 5 – 24.9									
No	3,566,615 <sup>c</sup>	15,170	3,581,785	329 (313, 339)	58.3	43.2	99.8	216.0	
Yes	8,243	11,532	19,775						
Modified height									
z-score of 3 – 24.9									
No	3,530,569	30,811	3,561,380	29.1 (28, 30)	20.2	20.9	99.1	23.2	
Yes	32,052	8,128	40,180						
Modified BMI									
z-score of 5 – 24.9									
No	3,550,256	25,370	3,575,626	254 (247, 263)	64.4	39.7	99.7	132.3	
Yes	9229	16,705	25,934						

<sup>a</sup>Children with low BIVs or modified z-score >25 (at either examination) are excluded.

<sup>b</sup>Odds ratio represents the odds of having a high BIV at the 2<sup>nd</sup> examination among those who had a high BIV at the 1<sup>st</sup> examination vs. the odds of a high follow-up BIV among those who did not have a high initial BIV

<sup>c</sup>Values are Ns for the cross-classification of BIVs at ages 2 y and 4 y

**Table 5**

Prevalence of extreme obesity based on current and expanded cut-points for the body size measures

Sex	Race	N	Current BIV cut points (+5 for Weight and BMI; +3 for Height) <sup>a</sup>			Z-scores of +8 and +4 (+227,667)			Z-scores of +15 and +5 (+264,353)			NHANES maximums (+232,316) <sup>b</sup>			No upper cut-point (+284,600)	
			N	Obesity	Severe Obesity	Obesity	Severe Obesity	Obesity	Severe Obesity	Obesity	Severe Obesity	Obesity	Severe Obesity	Obesity	Severe Obesity	Obesity
Overall		13,313,952	14.4	1.18	15.3	2.09	15.5	2.25	15.4	2.18	15.5	2.34	15.5	2.34	15.5	2.34
Boys		6,727,616	15.1	0.93	16.2	2.03	16.4	2.25	16.3	2.19	16.4	2.35	16.4	2.35	16.4	2.35
	White	2,261,561	13.0	0.64	13.7	1.42	13.9	1.59	13.8	1.53	13.9	1.63	13.9	1.63	13.9	1.63
	Black	1,312,383	12.2	0.61	13.0	1.33	13.2	1.51	13.1	1.45	13.2	1.57	13.2	1.57	13.2	1.57
	Hispanic	2,641,380	18.5	1.34	19.9	2.93	20.1	3.21	20.1	3.13	20.3	3.35	20.3	3.35	20.3	3.35
	Asian & Pacific Islander	178,258	12.9	0.87	13.8	1.77	13.9	1.91	13.8	1.87	14.0	1.98	14.0	1.98	14.0	1.98
	American Indian & Alaskan Native	51,478	21.2	1.23	22.6	2.77	23.0	3.20	22.8	3.05	23.0	3.24	23.0	3.24	23.0	3.24
Girls		6,586,336	13.8	1.44	14.5	2.15	14.6	2.24	14.5	2.18	14.6	2.33	14.6	2.33	14.6	2.33
	White	2,181,711	11.5	1.07	12.1	1.61	12.2	1.68	12.1	1.63	12.2	1.72	12.2	1.72	12.2	1.72
	Black	1,280,830	11.2	1.13	12.0	1.76	12.1	1.86	12.0	1.80	12.1	1.93	12.1	1.93	12.1	1.93
	Hispanic	2,625,843	17.0	1.94	17.9	2.83	18.0	2.93	17.9	2.86	18.1	3.06	18.1	3.06	18.1	3.06
	Asian & Pacific Islander	173,607	10.3	0.91	10.7	1.32	10.8	1.37	10.7	1.33	10.8	1.43	10.8	1.43	10.8	1.43
	American Indian & Alaskan Native	50,946	18.8	1.96	19.8	3.02	20.0	3.16	19.9	3.06	20.0	3.22	20.0	3.22	20.0	3.22

<sup>a</sup>Column headings show BIV cut-points based on modified z-scores of weight, BMI and height. The Ns show the increase in sample size as compared to the current cut-points in the CDC growth charts.

<sup>b</sup>The age-specific maximum values in NHANES 1999–2000 through 2011–12 for weight were 28 kg (age 2 y), 37 kg (3 y) and 50 kg (4 y); for height, the maximums were 105 cm (2 y), 114 cm (3 y) and 135 cm (4 y); for BMI, the maximums were 29.5 kg/m<sup>2</sup> (2 y), 33.2 kg/m<sup>2</sup> (3 y) and 33.0 kg/m<sup>2</sup> (4 y). These values are also shown in reference #22.