



Published in final edited form as:

*Ann Hum Biol.* 2017 December ; 44(8): 687–692. doi:10.1080/03014460.2017.1388845.

## Longitudinal changes in BMI z-scores among 45 414 2–4-year olds with severe obesity

David S. Freedman<sup>a</sup>, Nancy F. Butte<sup>b</sup>, Elsie M. Taveras<sup>c</sup>, Alyson B. Goodman<sup>a</sup>, and Heidi M. Blanck<sup>a</sup>

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

<sup>b</sup>Baylor College of Medicine, Children's Nutrition Research Center, Houston, TX, USA

<sup>c</sup>Department of Pediatrics, MassGeneral Hospital for Children, Boston, MA, USA

### Abstract

**Background**—BMI z-scores (BMIz) based on the Centers for Disease Control and Prevention (CDC) growth charts among children do not accurately characterise BMI levels among children with very high BMIs. These limitations may be particularly relevant in longitudinal and intervention studies, as the large changes in the L (normality) and S (dispersion) parameters with age can influence BMIz.

**Aim**—To compare longitudinal changes in BMIz with BMI expressed as a percentage of the 95th percentile (%BMI<sub>p95</sub>) and a modified z-score calculated as  $\log(\text{BMI}/M)/S$ .

**Subjects and methods**—A total of 45 414 2–4-year-olds with severe obesity (%BMI<sub>p95</sub> 120).

**Results**—Changes in very high BMIz levels differed from the other metrics. Among severely obese 2-year-old girls, for example, the mean BMIz decreased by 0.6 SD between examinations, but there were only small changes in BMI<sub>p95</sub> and modified BMIz. Some 2-year-old girls had BMIz decreases of >1 SD, even though they had large increases in BMI, %BMI<sub>p95</sub> and modified BMIz.

**Conclusions**—Among children with severe obesity, BMIz changes may be due to differences in the transformations used to estimate levels of BMIz rather than to changes in body size. The BMIs of these children could be expressed relative to the 95th percentile or as modified z-scores.

### Keywords

BMI; longitudinal; children; LMS

---

CONTACT: David S. Freedman [dxfl@cdc.gov](mailto:dxfl@cdc.gov) CDC F-77, 4770 Buford Highway, Atlanta, GA 30341–3724, USA.

#### Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

This work was authored as part of the Contributor's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

## Introduction

The 2000 Centers for Disease Control and Prevention (CDC) growth charts (Kuczmarski et al., 2002) are widely used to classify obesity (BMI ≥ 95th percentile for a child's sex and age) among children and adolescents in the US. These charts contain 10 smoothed percentiles (between the 3rd and 97th) of BMI (Flegal & Cole, 2013; Kuczmarski et al., 2002). These percentiles were used to derive three parameters (L, power transformation for normality; M, median; and S, dispersion) (Cole & Green, 1992) that allow the BMI of a child to be expressed as a  $z$ -score or percentile.

However, the use of these LMS parameters for BMIs above the CDC 97th percentile ( $z \sim 1.88$ ) constrains the maximum  $z$ -score that can be attained at a given sex and age (Centers for Disease Control and Prevention (CDC), 2016; Woo, 2009). A wide range of very high BMIs can, therefore, map to similar  $z$ -scores, depending upon the sex and age of the child (Freedman et al., 2017b). To overcome these limitations, it has been suggested that very high BMIs be expressed as a percentage of the CDC 95th BMI percentile (%BMI<sub>p95</sub>) (Flegal et al., 2009). Severe obesity is classified as a %BMI<sub>p95</sub> ≥ 120 (Kelly et al., 2013).

Although investigators have emphasised the drawbacks of analysing changes in BMI<sub>z</sub> (Berkey & Colditz, 2007; Cole et al., 2005; Kakinami et al., 2014; Paluch et al., 2007), these changes continue to be widely used in longitudinal studies (Lo et al., 2014; McCormick et al., 2014; O'Connor et al., 2017; Smego et al., 2017) and it has been suggested that a BMI  $z$ -score reduction of more than 0.20–0.25 SDs may be clinically important (US Preventive Services Task Force, 2017). The limitations of the CDC BMI<sub>z</sub> values may be particularly relevant in longitudinal studies of children with very high BMIs, as the age-related changes in the L and S parameters could strongly influence observed changes in BMI<sub>z</sub>. BMI<sub>z</sub> changes among these children may not accurately reflect changes in body size and analyses focused on BMI<sub>z</sub> change could lead to incorrect conclusions.

The objective of the current study is to examine longitudinal changes in BMI<sub>z</sub>, %BMI<sub>p95</sub> and a modified  $z$ -score that eliminates the compression of very high BMIs among 45 414 2–4-year-olds with severe obesity.

## Subjects and methods

As previously described (Freedman et al., 2016, 2017b), the CDC's Paediatric and Nutrition Surveillance Study (PedNSS) was a state-based public health surveillance system that monitored the nutritional status from birth to age 5 years of primarily low-income children. All procedures followed were in accordance with the ethical standards of CDC and approval was obtained from the relevant committee on human subjects.

## Data collection and statistical analyses

Weight was recorded to the nearest 1/4 pound and height to the nearest 1/8 inch with a measuring board (Mei et al., 1998). There were 14.4 million records from children who were between 24.0–59.9 months of age from 2008–2011. Based on the state, assigned ID and birthdate, we identified children who were examined multiple times.

The initial data cleaning excluded (1) 623 000 (4%) records that were missing information on weight or height (or had a value of 0), (2) 70 000 children for whom the date of birth or sex differed across examinations for the same ID and (3) 151 000 records that had a weight, height or BMI considered to be implausible based on levels of modified z-scores for weight, height, and BMI (Centers for Disease Control and Prevention (CDC), 2016).

These exclusions resulted in a dataset containing 13.5 million records from 8.7 million children. Of these children, 3.6 million were examined more than once and we focus on the first two examinations for the 3 570 241 children whose examinations occurred at least 3 months apart. We further excluded 51 114 children (1.4%) whose weight and height did not change between examinations, possibly indicating a data error.

There were also several children who appeared to have unreasonable changes in weight, height or BMI. For example, 0.1% of the children had annualised decreases of more than 3.8 kg for weight, more than 10 cm for height, and more than 10 kg/m<sup>2</sup> for BMI. We therefore excluded children who were in the upper or lower 0.1% of the distribution of these characteristics. These exclusions resulted in a sample of 3 498 745 children,

Of these children, 45 814 had a BMI  $\geq$  120% of the 95th percentile at their first examination and form the sample for the analyses. The mean interval between examinations was 11 (SD = 5) months.

### BMI metrics

Body mass index (BMI) was calculated as kg/m<sup>2</sup>. BMI<sub>z</sub> was calculated by expressing a child's BMI relative to children of the same sex and age in the CDC growth charts (Kuczmarski et al., 2002) using the L (transformation for normality) M (median) S (dispersion) method (Cole, 1990; Cole et al., 1998):

$$BMI_z = \left[ \left( \frac{BMI}{M} \right)^L - 1 \right] \div (L \times S)$$

Because L is less than -1 at all ages (the minimum is -3.4) (CDC; National Center for Health Statistics, n.d.), when BMI is large relative to M (median),  $(BMI \div M)^L$  approaches 0 and the maximum possible BMI<sub>z</sub> is  $-1 \div (L \times S)$ . For example, the value of  $(BMI \div M)^L$  for a 4-year-old girl with a BMI of 30 would be ~0.1, resulting in a BMI<sub>z</sub> of 3.4 SDs; the maximum possible BMI<sub>z</sub> for that sex/age, irrespective of the actual BMI, is 3.8 SDs. Between 24–59 months of age, the maximum possible BMI<sub>z</sub> in the CDC growth charts decreases from 11.9 to 3.5 SDs among girls, while the pattern among boys is concave, with the peak (~10 SDs) at about 39 months of age (Freedman et al., 2017b).

Obesity is defined as a BMI greater than or equal to the sex- and age-specific 95th percentile (BMI<sub>z</sub>  $\geq$  1.645) of the CDC growth charts (Kuczmarski et al., 2002; Ogden & Flegal 2010) and severe obesity as a BMI  $\geq$  120% of the 95th percentile (Kelly et al., 2013). We refer to a BMI that is expressed as a percentage of the 95th percentile as %BMI<sub>p95</sub>. For example, the

CDC 95th percentile for a 3-year-old boy is 19.3 kg/m<sup>2</sup> and a boy with a BMI of 21.3 kg/m<sup>2</sup> would have a %BMIp95 of 110% ( $100 \times 21.3/19.3$ ).

Although the L parameter in the CDC growth charts varies between -1.0 and -3.4, setting L equal to 0 removes the compression of very high BMIs into a narrow, bounded range of z-scores. Based on the Box-Cox transformation (Box & Cox, 1964), the equation for BMIz when L = 0 is:

$$\log\left(\frac{BMI}{M}\right) \div S$$

This modified z-score expresses a child's BMI relative to the median BMI and adjusts the log of this ratio for the dispersion (S) of the BMI distribution. It should be noted that this metric is unrelated to the method (Centers for Disease Control and Prevention (CDC), n.d.) that has been proposed to identify outliers in the CDC growth charts.

### Statistical analyses

Analyses were performed in R (R Core Team, 2017). We examine the cross-classification of BMIz and %BMIp95 between the two examinations and the changes by sex and both age and BMIz. These changes are displayed using violin plots (Hintze & Nelson, 1998), a technique that shows the entire distribution of the data rather than only selected percentiles, as do boxplots.

The intra-class correlation coefficient (ICC) (Shrout & Fleiss, 1979; Wolak, 2017) was used to assess the strength of the association across the two examinations. In contrast to the typical Pearson correlation coefficient, the ICC assesses the within-person similarity of the characteristic across examinations and its magnitude would be reduced if the mean levels differ between examinations. Standard errors of differences between the ICCs were obtained by bootstrapping (Canty & Ripley, 2017)

We also used time series plots to display changes in levels of the BMI metrics among the 100 girls who had an initial BMIz = 4.75; these line plots connect the initial and final values of BMIz (y-axis) for each child.

## Results

Table 1 shows levels of various characteristics at the first examination for the 45 414 children who had severe obesity at their first examination. Among these children, mean levels of BMIz were 3.9 SD (boys) and 3.4 SD (girls), and were 128–129 for %BMIp95. These children with severe obesity also had very high levels of weight-for-age and moderately elevated levels of height-for-age. The bottom of Table 1 shows the mean changes between examinations which occurred, on average, 11 months apart. Mean levels of all BMI metrics decreased with age, but, relative to the initial level, the mean changes in %BMIp95 and modified BMIz were very small. In contrast, the mean BMIz level among girls decreased by 0.5 SDs (~15% of the initial level).

Table 2 shows the cross-classification of BMIz (left) and %BMIp95 (right) levels at the two examinations. If BMIz levels were stable over time, most children would fall along the main diagonal, and roughly similar numbers would have lower or higher levels at re-examination. In contrast, most children had a substantially lower BMIz value upon re-examination and this was particularly evident among girls. Among the 8745 girls who had an initial BMIz of 3.5–4.4 SDs, for example, 73% had a BMIz of less than 3.5, while only 0.1% ( $n = 9$ ) had a higher BMIz. As assessed by the intra-class correlation coefficient, the association between BMIz levels at the two examinations among girls was very weak (ICC = 0.14, 95% CI = 0.13–0.15).

Substantially different results were seen for %BMIp95 (Table 2, right). For this analysis, cut points of %BMIp95 were selected so the sample size of each group at the initial examination would equal those for the BMIz categories. About 50% of girls with an initial %BMIp95 155 (vs 3% of girls in the highest BMIz category) remained in this category at re-examination (bottom row). The results for modified BMIz (not shown) were very similar to those for %BMIp95, with 50% of the girls in the highest initial category ( 7.5 SDs) remaining in this category at re-examination. As assessed by bootstrapping, the ICCs for %BMIp95 (0.46) and modified BMIz (0.48) at the two examinations were significantly ( $p < 0.0001$ ) higher than the ICC for BMIz (0.14) among girls.

Figure 1 shows violin plots, categorised by sex and age, for the changes in BMIz (top) and %BMIp95 (bottom) between examinations. Among 2-year-olds, the median BMIz among girls decreased by  $-0.6$  to  $-1.0$  SDs in the three categories of BMIz, while comparable changes among boys were much smaller (range =  $-0.2$  to  $+0.3$  SDs). The sex difference in BMIz changes was particularly striking for 331 2-year-olds who had an initial BMIz 4.5 SDs; 99% (328) of these girls showed a decrease (mean =  $-1.0$  SD) in BMIz between examinations. In contrast to BMIz, mean changes in %BMIp95 (bottom) showed little difference between boys and girls. The mean change in modified BMIz among 2-year-old girls with a BMIz 4.5 was  $+0.01$  SDs (not shown).

Figure 2 shows line plots for the 100 girls who had the highest initial value of BMIz ( 4.75), with each line connecting the initial and final value of the BMI metrics; age is on the  $x$ -axis. As seen in the left panel, the BMIz level of all girls decreased (mean =  $-1.1$  SDs) between examinations. Although most of these girls also showed a decrease in %BMIp95 between examinations, 35 showed an increase in %BMIp95 and 55 showed an increase in modified BMIz. Among these 100 girls, the median BMI decreased by  $1.7$  kg/m<sup>2</sup>, but the range of BMI changes varied from  $-6.3$  kg/m<sup>2</sup> to  $+2.6$  kg/m<sup>2</sup>.

Nine of these girls had a BMIz decrease by more than 1 SD, but a %BMIp95 increase of 10 or more percentage points. One girl, for example, had a BMIz decrease from 4.9 to 3.4 SDs between 26 and 55 months of age, but increases in BMI (from 28.0 to 30.6 kg/m<sup>2</sup>), %BMIp95 (from 148 to 169) and modified BMIz (from 6.5 to 8.6 SDs). Even if this girl's BMI had increased to 100 kg/m<sup>2</sup>, her BMIz would have decreased, because the maximum attainable BMIz in the CDC growth charts for a 55-month-old girl is 3.7 SDs.

## Discussion

Although the report for the CDC growth charts (Kuczmarski et al., 2002) cautioned about extrapolating outside the 3rd and 97th percentiles of BMI, BMIz remains widely used in studies that include large numbers of children with very high BMIs (Hampl et al., 2016; Kreier et al., 2013; McCormick et al., 2014; Smego et al., 2017; Wang et al., 2015). We have shown that, among children with severe obesity, BMIz levels are only weakly associated with other measures of body size (Freedman et al., 2017a) and the current results highlight the drawbacks of attempting to assess longitudinal BMIz changes. Because *z*-scores in the CDC growth charts for very high BMIs are strongly influenced by the sex- and age-specific values of the L and S parameters, an analysis of BMIz changes among children with severe obesity can lead to paradoxical results. For example, some children in the current study had substantial (  $\approx 1$  SD) decreases in BMIz, but large increases in BMI, %BMIp95 and modified BMIz. This is the result of differences in the L and S parameters at the initial and final examinations. We have found that the longitudinal tracking of BMI levels among 247 school-aged children with severe obesity is stronger for %BMIp95 than for BMIz ( $r = 0.61$  vs  $0.46$ ) (Freedman & Berenson, 2017).

We found, for example, particularly large decreases in the BMIz levels of girls who had very high BMIs at their initial examinations. This was the result of the maximum attainable BMIz among girls decreasing from 11.9 SDs (24 months) to 3.6 SDs (59 months), compressing very high BMIs into a lower range of *z*-scores at older ages. Although these decreases were most evident among girls with the very highest BMIz values, this compression affects a much larger group of children. For example, of the girls with severe obesity in the current study, only 13% showed an increase in BMIz between examinations, whereas 49% showed an increase in %BMIp95 and 38% showed an increase in modified BMIz.

The compression of very high BMI values into a narrow range of *z*-scores that differ by sex and age could be particularly problematic in intervention studies. Although the current study focused on 2–4-year-olds, a period in which the L parameter change markedly with age, the maximum attainable BMIz varies by sex and across all ages (Freedman et al., 2017a). Because of these differences in L and S by age, an ineffective intervention could result in a substantial decrease in BMIz or an effective treatment could result in no change depending on the sex and age. To alleviate the effects of the compression of BMIz, we used %BMIp95 and modified BMIz in the current study. The latter metric, which was based on setting the L parameter equal to 0, also accounts for differences in the dispersion of BMI values between boys and girls and across ages. We found similar results for both %BMIp95 and modified BMIz. It should be noted that other powers of L, such as 0.5, could also be used in the construction of modified *z*-scores.

It has been suggested that a BMI *z*-score reduction of more than 0.20–0.25 may be clinically important (US Preventive Services Task Force, 2017) and at least one intervention study has used a *z*-score decrease of  $\approx 0.2$  SD to classify a successful intervention (Markert et al., 2014). However, this criterion, particularly among very obese children, is problematic, because the interpretation of a 0.2 SD depends upon the initial BMIz value. Among 2-year-old girls in the current study, for example, 87% of those with an initial %BMIp95  $\approx 130$

showed a BMIz decrease of more than 0.2 SDs in the absence of any intervention. Among older children, in contrast, even large changes in BMI can be associated with virtually no change in BMIz. Based on the LMS parameters in the CDC growth charts, for example, the BMIz of a 94-month-old boy whose BMI decreased from 38 to 32 kg/m<sup>2</sup> over 1 year would decrease by only 0.16 SDs.

Many of the limitations of BMIz have been noted by others (Berkey & Colditz, 2007; Cole et al., 2005; Flegal et al., 2009; Paluch et al., 2007; Woo, 2009). Although the compression of very high values of BMIz occur with all growth charts, the effects are particularly large in the CDC growth charts, because BMIs in the CDC reference population (1960–1994) were more skewed and required larger, negative power transformations (L) (Woo, 2009) than in other growth charts. For example, the values of the L parameter in the WHO charts for both pre-school children (WHO Multicentre Growth Reference Study Group, 2006) and 5–19-year-olds (World Health Organization, n.d.) are also substantially closer to 0 than in the CDC growth charts. Because the WHO charts also estimate z-scores for values that are more than 3 SDs from the median by extrapolating the distance between 2–3 SDs, these growth standards do not show compression of very high BMIs.

Several limitations of our analyses should be considered. The PedNSS dataset is large, but the data were not collected for research purposes and there were many errors in the data. In addition to excluding a large number of records that were missing information on weight or height, we excluded children who had information for date of birth or sex that differed across examinations or who had a weight, height or BMI considered to be biologically implausible (Centers for Disease Control and Prevention (CDC), 2016; Freedman et al., 2016). We also excluded children from the analyses if their change in weight, height or BMI was in the upper or lower 0.1% of the distribution. Although it is likely that we did not exclude all errors and that some valid values were mistakenly excluded, any errors would be expected to have influenced changes in all BMI metrics. Finally, it should be realised that the longitudinal changes we observed in levels of BMIz, %BMIp95 and modified BMIz were based on cross-sectional growth charts and there are several problems with this approach (Cole, 1994; Zemel, 2009). For example, regression to the mean would result in subsequent values being closer to the median than is the initial level. Longitudinal velocity growth charts have been constructed to assess changes over time, but these have focused on weight and length among infants (Cole, 1995; WHO Department of Nutrition for Health and Development, 2009).

Extrapolated z-scores based on the CDC growth charts are widely used among children who have very high BMIs (Baughcum et al., 2015; Hampl et al., 2016; Kreier et al., 2013; Smego et al., 2017; Wang et al., 2015). However, BMIz values for extremely high BMIs can differ substantially from the empirical estimates (Flegal et al., 2009), have an effective upper limit (Woo, 2009) and are strongly influenced by sex and age. Our results indicate that the analyses of BMIz may be particularly problematic in longitudinal analyses of children with severe obesity. Among these children, changes in BMIz can simply reflect changes in various parameters of the CDC growth charts rather than changes in body size. In longitudinal studies of children who have very high BMIs, investigators should consider

using another BMI metric, such as %BMI<sub>p95</sub> or modified BMI<sub>z</sub> ( $\log(\text{BMI}/M) \div S$ ), rather than BMI<sub>z</sub>.

## Acknowledgments

### Funding

NB and ET were supported by funds from the Centers for Disease Control and Prevention [RFA-DP-11-007]. ET was supported by funds from the Centers for Disease Control and Prevention [U18DP003370] and by [grant K24 DK10589] from the National Institute of Diabetes and Digestive and Kidney Diseases.

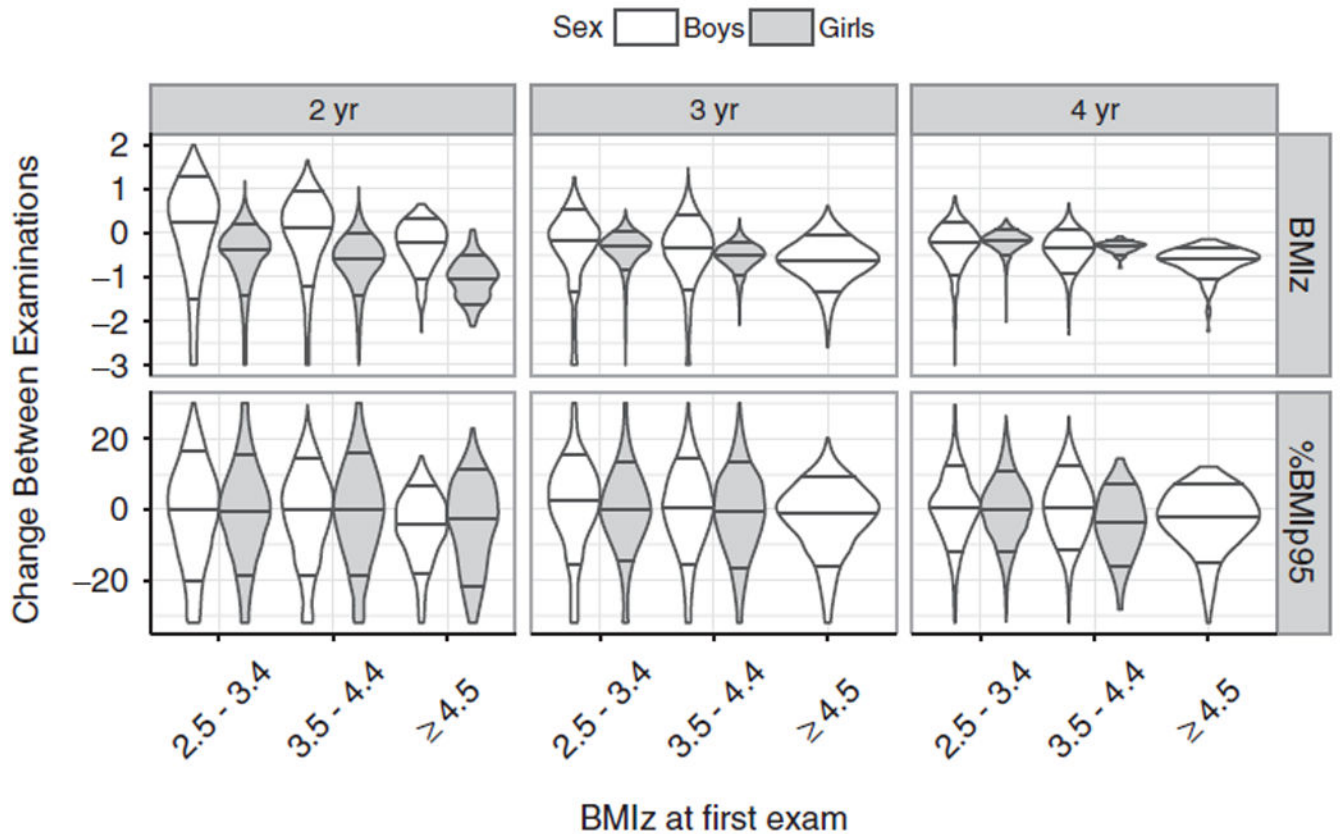
## References

- Baughcum AE, Gramling K, Eneli I. Severely obese preschoolers in a tertiary care obesity program: characteristics and management. *Clin Pediatr*. 2015; 54:346–352.
- Berkey CS, Colditz GA. Adiposity in adolescents: change in actual BMI works better than change in BMI z score for longitudinal studies. *Ann Epidemiol*. 2007; 17:44–50. [PubMed: 17140812]
- Box GEP, Cox DR. An analysis of transformations. *J R Stat Soc Series B*. 1964; 26:211–252.
- Canty, A., Ripley, B. *Boot: Bootstrap Functions* (Originally by Angelo Canty for S). 2017. Available at: <https://cran.r-project.org/package=boot>
- CDC. [May 10, 2017] National Center for Health Statistics, Growth Charts - Z-score Data Files. Available at: <https://www.cdc.gov/growthcharts/zscore.htm>
- Centers for Disease Control and Prevention (CDC). A SAS Program for the 2000 CDC Growth Charts. 2016. Available at: <https://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm>
- Centers for Disease Control and Prevention (CDC). [April 25, 2017] Modified z-scores in the CDC growth charts. nd. Available at: <https://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/biv-cutoffs.pdf>
- Cole TJ. Conditional reference charts to assess weight gain in British infants. *Arch Dis Child*. 1995; 73:8–16. [PubMed: 7639558]
- Cole TJ. Growth charts for both cross-sectional and longitudinal data. *Stat Med*. 1994; 13:2477–2492. [PubMed: 7701148]
- Cole TJ. The LMS method for constructing normalized growth standards. *Eur J Clin Nutr*. 1990; 44:45–60.
- Cole TJ, Faith MS, Pietrobelli A, Heo M. What is the best measure of adiposity change in growing children: BMI, BMI%, BMI z-score or BMI centile? *Eur J Clin Nutr*. 2005; 59:419–425. [PubMed: 15674315]
- Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. *Stat Med*. 1998; 17:407–429. [PubMed: 9496720]
- Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med*. 1992; 11:1305–1319. [PubMed: 1518992]
- 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]
- 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.
- Freedman DS, Butte NF, Taveras EM, Lundeen EA, Blanck HM, Goodman AB, Ogden CL. BMI z-Scores are a poor indicator of adiposity among 2- to 19-year-olds with very high BMIs, NHANES 1999-2000 to 2013-2014. *Obesity* (Silver Spring, Md). 2017a; 25:739–746.
- Freedman DS, Butte NF, Taveras EM, Goodman AB, Ogden CL, Blanck HM. The limitations of transforming very high body mass indexes into z-scores among 8.7 million 2- to 4-year-old children. *J Pediatr*. 2017b; 188:50–56.e1. [PubMed: 28433203]



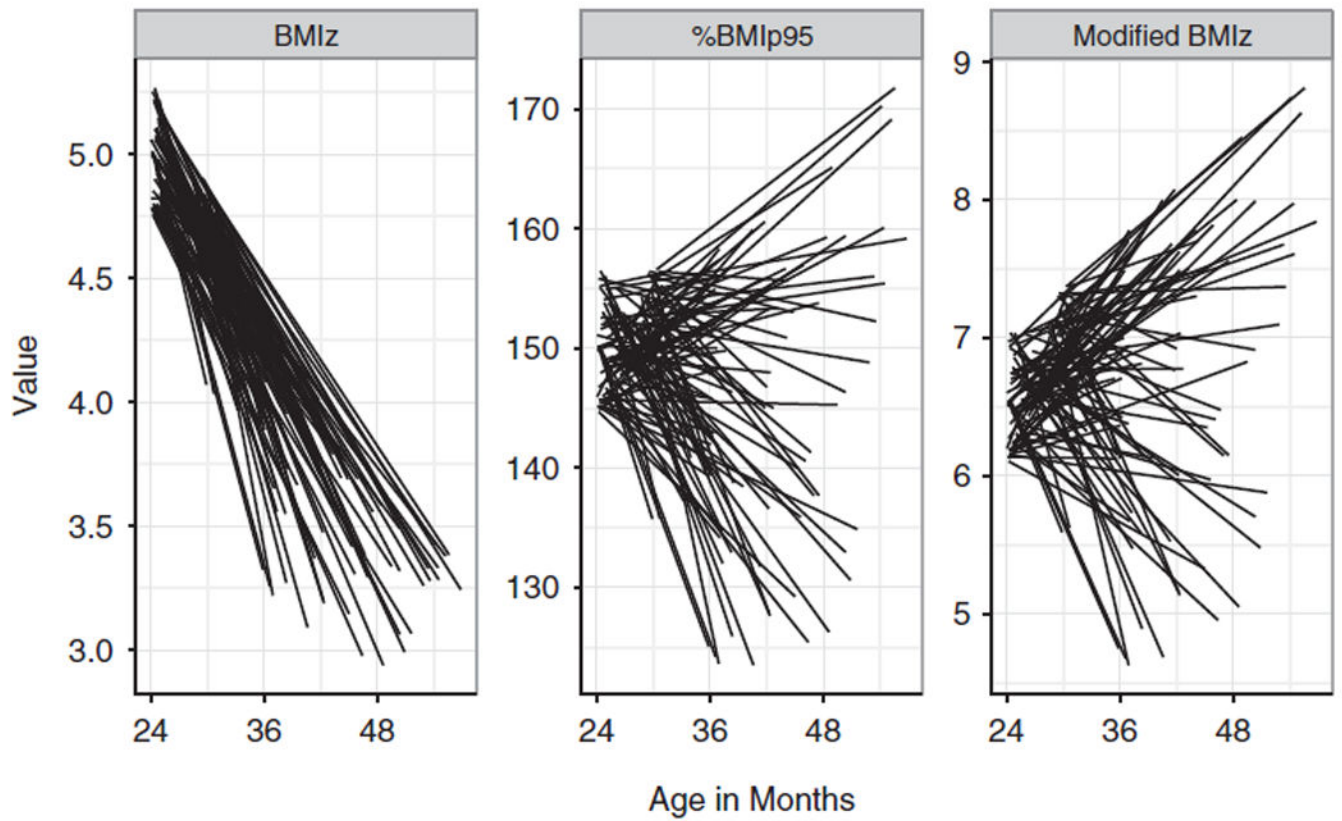
- Freedman DS, Lawman HG, Pan L, Skinner AC, Allison DB, McGuire LC, Blanck HM. The prevalence and validity of high, biologically implausible values of weight, height, and BMI among 8.8 million children. *Obesity* (Silver Spring, Md). 2016; 24:1132–1139.
- Freedman DS, Berenson GS. The longitudinal tracking of BMI z-scores and other BMI metrics among 2- to 17-year-olds. *Pediatrics*. 2017; In Press. doi: 10.1542/peds.2017-1072
- HAMPL S, Odar Stough C, Poppert Cordts K, Best C, Blackburn K, Dreyer Gillette ML. Effectiveness of a hospital-based multidisciplinary pediatric weight management program: Two-year outcomes of PHIT Kids. *Child Obes*. 2016; 12:20–25. [PubMed: 26790094]
- Hintze JL, Nelson RD. Violin plots: a box plot-density trace synergism. *Am Stat*. 1998; 52:181–184.
- Kakinami L, Henderson M, Chiolero A, Cole TJ, Paradis G. Identifying the best body mass index metric to assess adiposity change in children. *Arch Dis Child*. 2014; 99:1020–1024. [PubMed: 24842797]
- Kelly AS, Barlow SE, Rao G, Inge TH, Hayman LL, Steinberger J, Urbina EM, et al. Severe obesity in children and adolescents: identification, associated health risks, and treatment approaches: a Scientific Statement from the American Heart Association. *Circulation*. 2013; 128:1689–1712. [PubMed: 24016455]
- Kreier F, Genco M, Boreel M, Langkemper MP, Nugteren IC, Rijnveld V, Thissen V, et al. An individual, community-based treatment for obese children and their families: the solution-focused approach. *Obesity Facts*. 2013; 6:424–432. [PubMed: 24107796]
- Kuczumski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, Wei R, et al. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat*. 2002; 11:1–190.
- Lo JC, Chandra M, Sinaiko A, Daniels SR, Prineas RJ, Maring B, Parker ED, et al. Severe obesity in children: prevalence, persistence and relation to hypertension. *Int J Pediatr Endocrinol*. 2014; 2014:3. [PubMed: 24580759]
- Markert J, Herget S, Petroff D, Gausche R, Grimm A, Kiess W, Blüher S. Telephone-based adiposity prevention for families with overweight children (T.A.F.F.-Study): one year outcome of a randomized, controlled trial. *Int J Environ Res Public Health*. 2014; 11:10327–10344. [PubMed: 25286167]
- McCormick EV, Dickinson LM, Haemer MA, Knierim SD, Hambidge SJ, Davidson AJ. What can providers learn from childhood body mass index trajectories: a study of a large, safety-net clinical population. *Acad Pediatr*. 2014; 14:639–645. [PubMed: 25129568]
- 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.
- O'Connor EA, et al. Screening for Obesity and Intervention for Weight Management in Children and Adolescents. Evidence Report and Systematic Review for the US Preventive Services Task Force (USPSTF). *JAMA*. 2017; 317:2427. [PubMed: 28632873]
- Ogden CL, Flegal KM. Changes in terminology for childhood overweight and obesity. *Natl Health Stat Report*. 2010; 25:1–5.
- Paluch RA, Epstein LH, Roemmich JN. Comparison of methods to evaluate changes in relative body mass index in pediatric weight control. *Am J Hum Biol*. 2007; 19:487–494. [PubMed: 17546615]
- R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017. Available at: <http://www.r-project.org/>
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979; 86:420–428. [PubMed: 18839484]
- Smego A, Woo JG, Klein J, Suh C, Bansal D, Bliss S, Daniels SR, et al. High body mass index in infancy may predict severe obesity in early childhood. *J Pediatr*. 2017; 183:87–93.e1. [PubMed: 27916426]
- US Preventive Services Task Force. Screening for obesity in children and adolescents: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2017; 317:2417–2426. [PubMed: 28632874]

- Wang Y, Cai L, Wu Y, Wilson RF, Weston C, Fawole O, Bleich SN, et al. What childhood obesity prevention programmes work? A systematic review and meta-analysis. *Obes Rev.* 2015; 16:547–565. [PubMed: 25893796]
- WHO Department of Nutrition for Health and Development. [May 17, 2017] WHO child growth standards. Growth velocity based on weight, length and head circumference. Methods and development. 2009. Available at: [http://www.who.int/childgrowth/standards/velocity/tr3\\_velocity\\_report.pdf](http://www.who.int/childgrowth/standards/velocity/tr3_velocity_report.pdf)
- WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva: World Health Organization; 2006.
- Wolak, M. [February 22, 2017] ICC: Facilitating Estimation of the Intraclass Correlation Coefficient [R package ICC version 2.3.0]. 2017. Available at: <https://cran.r-project.org/package=ICC>
- Woo JG. Using body mass index Z-score among severely obese adolescents: a cautionary note. *Int J Pediatr Obes.* 2009; 4:405–410. [PubMed: 19922058]
- World Health Organization. Growth reference 5-19 years WHO | BMI-for-age (5-19 years). WHO; n.d. Available at: [http://www.who.int/growthref/who2007\\_bmi\\_for\\_age/en/](http://www.who.int/growthref/who2007_bmi_for_age/en/) [August 2, 2017]
- Zemel BS. A commentary on the construction of weight velocity charts. *Nutr Clin Pract.* 2009; 24:651–653. [PubMed: 19841252]



**Figure 1.**

Change in BMIz (top) and %BMIp95 (bottom) between the two examinations among children who were severely obese at their initial examination. Results are stratified by sex and initial year of age. The violin plots represent the density distribution of the observed changes within each BMIz group and the horizontal lines represent the 10th, 50th and 90th percentiles.



**Figure 2.** Line plots showing changes in BMIz and %BMIp95 among the 100 girls with the highest BMIz values ( 4.75) at the initial examination. Each line represents the value of the BMI metric at the two examinations for a single girl. Age is shown on the  $x$ -axis.

**Table 1**

Levels of various characteristics among the 45 414 children who had severe obesity at their first examinations.

	Boys	Girls
<i>n</i>	21 007	24 407
Age in months	37.4 (24.4, 51.9) <sup>a</sup>	36.9 (24.3, 51.5)
BMI (kg/m <sup>2</sup> )	23.5 (21.6, 26.6)	23.7 (21.8, 27.7)
BMIz (SDs)	3.9 (3.2, 4.8)	3.4 (2.8, 4.4)
%BMIp95	128.2 (120.2, 145.7)	129.3 (120.3, 151.3)
Modified BMIz (SDs)	5.2 (4.3, 7.0)	5.1 (4.1, 7.2)
Weight-for-age <i>z</i> -score	3.4 (1.3, 5.1)	3.1 (1.4, 4.4)
Height-for age <i>z</i> -score	0.8 (−2.6, 3.1)	0.9 (−2.2, 3.2)
Changes in the BMI metrics <sup>b</sup>		
BMI (kg/m <sup>2</sup> )	−0.7 (−5.9, 3.2)	−0.5 (−5.6, 3.8)
BMIz (SD)	−0.3 (−2.3, 1.2)	−0.5 (−2.0, 0.3)
%BMIp95	−0.9 (−28.7, 20.7)	−0.8 (−27.9, 23.0)
Modified BMIz	−0.1 (−3.4, 2.2)	0.0 (−2.9, 2.3)

<sup>a</sup>Values in parentheses include 95% of the values (2.5 percentile to 97.5 percentile).

<sup>b</sup>Calculated as Exam 2 value – Exam 1 value. The mean interval between examinations varied was 10.8 months.

**Table 2** Cross-classification of BMIz levels at the first and second examinations among children who were severely obese ( $n = 45\ 414$ ) at their initial examination

	<i>n</i>	BMIz				%BMIp95 <sup>a</sup>			
		Initial category		Second examination		Initial category		Second examination	
		<3.5	3.5–4.4	3.5–4.4	4.5	Low	Middle	High	
Boys	4 202	<3.5	37%	11%	120–121	52%	40%	8%	
	14 690	3.5–4.4	38%	50%	122–138	31%	48%	20%	
	2 115	4.5	10%	66%	139	6%	35%	59%	
	15 331	<3.5	96%	4%	120–129	70%	29%	0.9%	
Girls	8 745	3.5–4.4	73%	27%	130–154	28%	63%	9%	
	331	4.5	39%	58%	155	2%	49%	50%	

<sup>a</sup>The sex-specific cut points for %BMIp95 at the first examination were chosen to make the number of children in each group equal to that in each BMIz group. Cut-points for %BMIp95 at the second examination correspond to those at the first examination: <122, 122–138 and 139 for boys, and <130, 130–154 and 155 for girls.

<sup>b</sup>Values represent the percentage of children in each initial category who are in the specified category at the second examination.