

# Challenges of Accurately Measuring and Using BMI and Other Indicators of Obesity in Children

**CONTRIBUTOR:** John H. Himes, PhD, MPH

*Division of Epidemiology and Community Health, University of Minnesota, School of Public Health, Minneapolis, Minnesota*

**KEY WORDS**

body mass index, overweight, obesity, child obesity, measurement

**ABBREVIATIONS**

CDC—Centers for Disease Control and Prevention

IOTF—International Obesity Taskforce

WHO—World Health Organization

CI—confidence interval

BIA—bioelectrical impedance analysis

[www.pediatrics.org/cgi/doi/10.1542/peds.2008-3586D](http://www.pediatrics.org/cgi/doi/10.1542/peds.2008-3586D)

doi:10.1542/peds.2008-3586D

Accepted for publication Apr 29, 2009

Address correspondence to John H. Himes, PhD, MPH, University of Minnesota, School of Public Health, Division of Epidemiology and Community Health, 1300 S 2nd St, Suite 300, Minneapolis, MN 55454. E-mail: [himes001@umn.edu](mailto:himes001@umn.edu)

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2009 by the American Academy of Pediatrics

**FINANCIAL DISCLOSURE:** *The author has indicated he has no financial relationships relevant to this article to disclose.*

## abstract

BMI is an important indicator of overweight and obesity in childhood and adolescence. When measurements are taken carefully and compared with appropriate growth charts and recommended cutoffs, BMI provides an excellent indicator of overweight and obesity that is sufficient for most clinical, screening, and surveillance purposes. Accurate measurements of height and weight require that adequate attention be given to data collection and management. Choosing appropriate equipment and measurement protocols and providing regular training and standardization of data collectors are critical aspects that apply to all settings in which BMI will be measured and used. Proxy measures for directly measured BMI, such as self-reports or parental reports of height and weight, are much less preferred and should only be used with caution and cognizance of the limitations, biases, and uncertainties attending these measures. There is little evidence that other measures of body fat such as skinfolds, waist circumference, or bioelectrical impedance are sufficiently practicable or provide appreciable added information to be used in the identification of children and adolescents who are overweight or obese. Consequently, for most clinical, school, or community settings these measures are not recommended for routine practice. These alternative measures of fatness remain important for research and perhaps in some specialized screening situations that include a specific focus on risk factors for cardiovascular or diabetic disease. *Pediatrics* 2009;124:S3–S22

BMI (weight [kg]/height [m<sup>2</sup>]) has probably become the most common indicator used to assess overweight and obesity in a wide variety of settings, including clinical, public health, and community-based programs. Although it is certainly not a perfect surrogate for total body fatness and not without its technical limitations,<sup>1</sup> BMI has been recommended as the most appropriate single indicator of overweight and obesity in children and adolescents outside of research settings.<sup>2-4</sup>

One of the attractive features of BMI is that it is derived from measurements of height and weight. These 2 anthropometric dimensions are the ones most commonly collected on children worldwide. These 2 measurements are noninvasive, relatively inexpensive to obtain, and relatively easily understood by health practitioners, the individuals being measured, and their families.

Mentioning child measurements of height and weight, individuals may be reminded of their own marks on the door sills and the bathroom scales of their childhood homes. So, although wide familiarity with height and weight enhances the use and understanding of a measure such as BMI, it also may desensitize health professionals to the need to give adequate attention to issues concerning how height and weight data are collected. Accordingly, one may hear the comment, "Anyone can measure height and weight." Although one must actually agree with the language, if not the intent, of this easy declaration, many health professionals are unaware that there are consequences for the usefulness and accurate interpretation of BMI data that follow from decisions made concerning data collection.

In this article, challenges surrounding the measurement of BMI in US children (2–18 years of age) and the implications of these issues for the

appropriate collection, use, and interpretation of BMI as it is used as an indicator of child and adolescent overweight and obesity are considered. Also, chief measurement issues related to other selected anthropometric indicators of overweight and obesity are briefly discussed.

### **SOME BASIC CONCEPTS FROM MEASUREMENT THEORY**

Classical measurement theory includes some concepts that are helpful for understanding issues surrounding measurement of height, weight, and, therefore, BMI. Detailed explanations of measurement theory are available in standard textbooks concerning measurement and psychometrics.<sup>5,6</sup> Different academic disciplines may use different terms to refer to the same concepts, but for the present discussion the terms usually found in the biomedical and epidemiologic literature will be used.

It is important to know that all measurements are imperfect and always measured with some error, whether the measurements be height, weight, skinfolds, or bioelectric impedance. Accordingly, an index such as BMI, which is derived from 2 other measurements, will include the components of measurement error inherent in the constituent height and weight measurements. The nature and magnitude of these measurement errors have some fairly predictable consequences related to the usefulness and interpretation of the measurements.

Some measurement errors are random, with the same probability of being smaller than or greater than the true value (a theoretical value measured without error). Consequently, the average or mean of random errors across a series of measurements is 0. For example, nurse Brown measured heights on a group of 4-year-old girls on Monday, and the mean height was

100 cm. She measured the same children a second time on Tuesday, again with a mean height of 100 cm. Nevertheless, for some girls there were small differences in height measurements between Monday and Tuesday, although the mean height of all girls remained the same for the 2 days. The differences between measured heights on Monday and Tuesday for the individual girls are examples of random errors of measurement.

Random errors of measurement are a concern, because they always add to the variability of the true measurements; their presence and extent are usually considered the measurement's "reliability." Poor measurement reliability is a concern because it may cause incorrect clinical judgments for individual children (misclassification) and alter conclusions for statistical analysis for groups of children. Because most inferential statistical tests use a measure of variation (eg, SD) as a denominator, statistical tests of differences between means, analysis of variance, correlations, regressions, and odds ratios are all attenuated (ie, less statistically significant) as the measurement reliability decreases and the variability term in the denominator increases. Random errors are usually reported in terms of a measurement error variance or a measurement error SD, or summarized in reliability coefficients (interclass or intraclass correlations) from replicate measurements of the same children.

In a second example, nurse Brown measured the same group of girls on Monday, again with a mean height of 100 cm. This time on Tuesday nurse Jones measured them for a second time and recorded heights exactly 1.0 cm taller than did nurse Brown for every girl. Now the mean height for all the girls on Tuesday was 101 cm. If we consider nurse Brown to be our gold

standard of measurement, this systematic measurement error (ie, all in 1 direction) of nurse Jones is an example of measurement bias.

Measurement bias is a concern because it may cause misclassification of individual children or groups of children. Nevertheless, as long as the bias is not differential among groups, pure measurement bias will not affect the results of statistical tests between or among groups, such as differences between means, analysis of variance, correlations (interclass), regressions, and odds ratios. In practice, differences between individual observers who measure the same children will also have a component of random measurement error between them. Not surprisingly, observers tend to measure more like themselves than like others, so interobserver errors are almost always larger than intraobserver errors.

Measurement theory usually specifies that measurement errors are independent and additive, that is, that the total measurement error variance is the sum of error variances from all sources.<sup>5</sup> Also, when increments or differences between successive measurements are used, the measurement errors attending each of the 2 constituent measurements are included with the increment. So an increment has twice the random measurement error (variance) of an attained value and lower measurement reliability. Obviously, if measurement biases change over time, the accuracy of increments becomes questionable.

### CHIEF SOURCES OF MEASUREMENT ERRORS FOR HEIGHT AND WEIGHT

When a child's height and weight are measured, there are several possible sources of measurement error. A simplified theoretical model would say that the total variance of measurement error is the sum of that

associated with the instrument used to measure, that associated with the child being measured, and that associated with the observer(s) doing the measuring. In most settings, however, errors associated with the child and with the observer(s) are the chief sources of measurement error in measurements of height and weight. Obviously, it is still important to have appropriate measuring equipment, but once they are installed and calibrated, little measurement error usually is due to the instruments per se.

### Measurement Errors Due to Child Variation

The normal day-to-day variation within a child leads to a component of measurement error. This variation probably results from many sources including hydration, gastrointestinal and urinary bladder contents, diurnal hormonal fluctuations, saltatory growth, fidgeting, alterations in position, and fatigue.<sup>7,8</sup>

As early as 1724, Wasse recognized appreciable variation in stature during the day and concluded that "[t]he alteration in the human stature . . . proceeds from the yielding of the cartilages between the vertebrae to the weight of the body in an erect posture."<sup>9</sup> MRI studies have since confirmed that the diurnal variation in stature primarily results from increases in water content in the soft central portion of the intervertebral discs (nucleus pulposus) while at rest and water loss while standing or during other weight-bearing activities.<sup>10</sup> For children, one can expect a mean height difference of ~1.5 cm (SD: 0.46 cm) between rising and late afternoon,<sup>11</sup> with most of the change probably occurring during the first 2 to 3 hours of the day.<sup>12</sup>

In practice, it is helpful to understand the expected diurnal variation in child height but probably impractical to try

to time height measurements to accommodate it, unless one is engaged in a rigorous research protocol that requires serial measurements on a small number of individual children. For the data included in the 2000 Centers for Disease Control and Prevention (CDC) growth charts,<sup>13</sup> heights were measured from mornings through evenings so that the reference percentiles represent something like heights averaged throughout the day, and the associated within-child variation is included in the total variance in height captured in the published percentiles or z scores at an age.

For body weight, the within-child variation is related to the size of the child and should usually be within 1.5% of the measured weight (SD: 0.5%).<sup>14</sup> Accordingly, the expected maximum within-child weight variation for children who weigh 25 and 50 kg should be ~375 g (0.83 lb) and 750 g (1.65 lb), respectively. In practice, it is difficult to standardize this physiologic within-child weight variation when children are measured, so it is usually ignored for most purposes.

### Seasonal Variation

It has been known for a long time that in some environments children may grow differentially according to season of the year.<sup>15</sup> It is important, however, to understand the contexts of these findings to determine the implications for current studies of height, weight, and BMI.

In developing countries with prevalent poverty, undernutrition, and infection, reduced seasonal patterns of average growth in height and weight are often linked to the rainy season(s), along with accompanying factors including reduced food availability and increased infection.<sup>16,17</sup> In developed countries the evidence is mixed, but when seasonal patterns are present, they usually indicate relatively greater

growth in height and linear dimensions during the spring and summer and relatively greater growth in weight and fatness during the fall and winter.<sup>18,19</sup> When seasonal fluctuations exist in developed countries, they are smaller and less common than those seen in children living in developing countries.

In studies in both Japan<sup>20</sup> and the United States,<sup>21</sup> seasonal fluctuations in growth were observed in earlier generations of children but disappeared within the same populations over 20 to 40 years because general health and nutrition conditions improved through time. Accordingly, for almost all children now living in the United States, there should be little if any seasonal variation in growth that would require accounting for it in the design of studies or data-collection protocols.

Excess growth in BMI has been observed over summer vacation between kindergarten and first grade for children in the Early Childhood Longitudinal Survey.<sup>22</sup> Nevertheless, this should probably be viewed as a school/no-school effect rather than seasonal variation per se.

### Measurement Errors Due to Observer Variation

An important goal in measurement of height and weight should always be to collect the data with as little measurement error as possible, given the practical and financial constraints of the local situation.

In a highly controlled research laboratory with experienced anthropometrists, the mean interobserver (absolute) differences for standing height and weight are 0.3 cm and 0.02 kg, respectively, with corresponding SDs of 0.2 cm and 0.03 kg.<sup>23</sup> These values should be viewed as close to the minimum values possible using current methods. In most situations there is

more concern about observer reliability in measurements of height rather than weight, because height measurements include more “opportunities” for within-child and observer variation than do weight measurements.

Often, height and weight measurements for BMI are collected in clinical or other settings in which data collection may be hurried and observers may not have been trained as rigorously as observers in research settings. Actually, there are few studies available concerning measurement variation among those who probably collect most of the data used for BMI evaluation and screening. Ahmed et al<sup>24</sup> evaluated the measurement variation among 2 sets of health visitors who each measured each of 10 children at ages 3 and 4.5 years 3 times with a portable stadiometer. The average value for the SD of measurement was 0.47 cm. In a small comparison trial on height of 5- and 6-year-old British children, school nurses had a pooled interobserver measurement SD of 0.32 cm, which compared favorably to that of a trained auxologist (0.35 cm) on the study.<sup>25</sup> The nurses in this study had been trained in measuring height. Importantly, training can improve the precision of length and height measurements.<sup>8,26</sup>

Given the above-listed principles, it follows that when a large number of data collectors are required the interobserver measurement errors increase as well.<sup>27</sup> Consequently, one would prefer to have as few individuals measuring height and weight as is practicable in the particular setting, especially if the resulting data will be used for research purposes or if serial measurements on the same children are being made.

Another strategy for reducing measurement errors is to take the measurements more than once and then use the mean of the replicates. The

theory here is that a mean of replicates is a better estimate of the “true” measurement, because the random errors of measurement are reduced.<sup>28</sup> The usefulness of taking replicate measurements depends on the reliability of the single measurement in question and how the data will be used.

Routinely obtaining replicates benefits most those measurements that have the lowest initial reliability, and the corresponding improvements in reliability are predictable.<sup>28</sup> Measurement-reliability coefficients ( $R$ ) express the percentage of the total observed variation that is captured by the “true” measurement variation. For single measurements of height and weight in a nonresearch setting, a reasonable expectation for values of  $R$  should be  $\sim 0.93$  and  $0.97$ , respectively. At these levels of measurement reliability, collecting a second measurement and using the mean raises the values of  $R$  to  $0.963$  and  $0.984$ , respectively. These are not dramatic improvements in measurement reliability using a duplicate, because the initial levels of measurement reliability started out rather high.

Contrast these possible improvements when using replicate measurements with those for skinfold thicknesses, for which the measurement reliability for a single measurement in nonresearch settings is probably  $\sim 0.8$ . For successive numbers of replicate skinfold measurements and using the mean, the  $R$  values would be  $0.88$  for 2 measurements,  $0.92$  for 3 measurements, and  $0.94$  for 4 measurements.

The errors of measurement with low measurement reliability are usually assumed to be largely random. Consequently, how the data are to be used is a consideration in deciding whether the extra time and trouble should be spent routinely collecting replicate measurements. Purely random errors



will not affect the group means of height, weight, and BMI, although they will increase the SDs because of the added error variance. Similarly, the prevalence of children with a BMI above percentile cutoffs for age and gender will not be affected by the added random error because as many children should be misclassified above and below the cutoff value. If the BMI data are to be used for these purposes, routinely taking replicate measurements is probably not worthwhile. For some uses of BMI data, however, routinely taking replicate measurements is recommended. If the BMI data will be used to make clinical decisions regarding treatment or referral of individual children, or for assessing changes in individuals over time, a second measurement of height and weight will reduce misclassification of current status and increase the ability to detect changes from one occasion to another. In research settings that include height, weight, and BMI as important variables, duplicate measurements of height and weight are recommended. If the height and weight replicates are averaged before calculating BMI, the latter calculation only needs to occur once.

## CHALLENGES OF USING APPROPRIATE REFERENCE DATA AND CUTOFFS

### Which Reference Data?

Usually, BMI will be evaluated in children relative to reference data or growth charts. The main challenge to the investigator is to choose the set of growth charts that is most appropriate for the intended purposes for which the BMI data will be used. For height, weight, and BMI, US investigators have the benefit of recent recommendations from an expert committee.<sup>4,29</sup>

For most purposes, US children aged 2 to 18 years should be evaluated

relative to the 2000 CDC growth charts.<sup>13</sup> These are high-quality growth charts that present selected percentiles and allow calculation of z scores of attained height, weight, and BMI for age and gender and in metric and English units. The primary data were collected in national surveys by using rigorous measurement protocols, and state-of-the-art statistical methods were used to derive and smooth the percentiles and z scores across the ages. More detailed technical information on methods and development are available elsewhere.<sup>13</sup> Earlier sets of BMI reference data for US children (eg, Must et al<sup>30</sup>) should not be used because the cutoff values are slightly different, which will serve to complicate comparisons across studies.

Some other countries have developed and use their own growth charts, but 2 sets designed for international applications should be briefly mentioned, particularly relative to BMI. The International Obesity Taskforce (IOTF) sponsored a workshop with a goal of establishing a standard definition for child overweight and obesity worldwide.<sup>31</sup> As a result, high-quality BMI data from 6 countries (Brazil, Great Britain, Hong Kong, Netherlands, Singapore, and the United States) were combined to develop age- and gender-specific cutoffs for children (birth to 20 years of age) corresponding to the locations of the BMI values of 25 and 30 kg/m<sup>2</sup> in the statistical distribution of adults.<sup>19</sup> These latter BMI cutoffs are the conventional criteria that identify overweight and obesity in adults.<sup>32</sup>

The IOTF cutoffs that define overweight and obesity correspond approximately to percentiles 82 to 84 and 96 to 97, respectively, on the 2000 CDC growth charts for BMI for age, not very different from the 85th- and 95th-percentile cutoffs used customarily in the United States. Prevalences of overweight and obesity in children in countries outside

the United States are now being reported in the literature rather frequently using the IOTF criteria,<sup>33,34</sup> which has been useful in standardizing BMI criteria. Nevertheless, it should be noted that the IOTF charts contain no percentile or z-score curves other than the 2 cutoff lines, because they were specifically designed for reporting population prevalences of overweight and obesity. Accordingly, the IOTF charts should not be used to monitor BMI growth in individual children.

In 2006 the Department of Nutrition and Health at the World Health Organization (WHO) released a new growth standard for children from birth to 5 years of age based on longitudinal and cross-sectional data collected in 6 countries (Brazil, Ghana, India, Norway, Oman, and United States).<sup>35</sup> The new attained growth curves, including BMI, were designed to represent how all children ought to grow under ideal circumstances. Accordingly, the mothers and children were carefully selected so that there were no known constraints to healthy growth, including exclusive breastfeeding and appropriate introduction of solid foods.<sup>36</sup> Because of the homogeneous nature of the WHO samples and some choices made to exclude the heaviest children, the upper BMI percentiles and z scores are somewhat restricted (ie, narrower) at an age compared with those in the 2000 CDC growth charts. Consequently, using the same percentile cutoff for BMI at an age (eg,  $\geq 95$ th), the WHO standards will yield a higher prevalence of children than if the  $\geq 95$ th percentile for age were used from the 2000 CDC growth charts.<sup>37</sup> The opposite is true at the other end of the BMI distribution so that thinness defined by a low BMI percentile on the WHO standards will identify fewer children with low BMI compared with using the same percentile cutoff on the 2000 CDC growth charts.<sup>38</sup>

One concern about using these new WHO growth standards is the interpretation in terms of the health or growth of children who are in the extremes of the percentiles (eg, <5th, >95th) on the basis of a standard that purportedly only included healthy children. Nevertheless, the WHO standards are so new that there are no data documenting whether the new cutoffs are better at identifying children at health risk than the 2000 CDC growth charts.

In 2007, the WHO released a growth reference for height, weight, and BMI for children aged 5 to 19 years that was designed to align with the 2006 WHO growth standards at 5 years and to be used internationally.<sup>39</sup> The WHO re-analyzed the data comprising the US National Center for Health Statistics growth curves, published in 1977,<sup>40</sup> and proposed that they be used as a single growth reference for screening, surveillance, and monitoring of school-aged children worldwide. As with children older than 24 months included in the new WHO birth to 5 years reference,<sup>35</sup> BMI values of >2 SDs were excluded as unhealthy for the 2007 5 to 19 years reference.<sup>39</sup> Because the heaviest children were excluded, the upper percentiles of BMI for the WHO 2007 reference are substantially below the corresponding levels for the 2000 CDC growth charts, especially in later adolescence when high BMI values are more common.

There has been much informal discussion about the use of the IOTF and WHO references. Unfortunately, there have been no formal recommendations from agencies or professional organizations in the United States regarding their routine or partial use (eg, at certain ages or for certain purposes). This institutional silence is unfortunate, because it will likely lead to at least ambiguity and perhaps even confusion among health practitioners and in the scientific literature.

As a personal recommendation for health practitioners in the United States, the 2000 CDC growth charts should be used for routine screening, surveillance, and monitoring of BMI because they have been widely evaluated and adopted, and they have been recommended by recent expert committees.<sup>4,29</sup> If investigators wish to communicate with international colleagues in presentations and in the scientific literature by citing the IOTF or WHO criteria, they should also include at least prevalence results relative to the 2000 CDC growth charts so that their findings can be compared with those of other US studies. Hopefully, as further research becomes available, more specific recommendations can be made on the basis of studies of sensitivity/specificity and differential risk among the various BMI criteria currently available.

### A Rose by Any Other Name

Before 1994 the scientific literature on overweight and obesity included a wide range of defining criteria (eg, percent ideal weight, skinfold thickness, ponderal index, BMI) and many descriptive names to refer to the children and adolescents who were considered the fattest. This variation in reporting made it difficult to compare findings because different indicators may actually identify different children as the fattest,<sup>41</sup> and the differences in terminology were sometimes confusing. An expert committee considered these issues, and their proceedings, published in 1994,<sup>2</sup> had considerable effect toward standardizing the criteria (BMI for age) and the nomenclature for referring to the fattest children and adolescents. Subsequently, these definitions became preferred in describing weight status.<sup>3,42,43</sup>

In the 1994 report,<sup>2</sup> children with a BMI that exceeded 30 kg/m<sup>2</sup> or ≥95th percentile for age and gender (whichever

was smaller) were considered overweight. Children or adolescents with a BMI at ≥85th percentile but <95th percentile were considered at risk of overweight. At that time, the term “obese” was avoided, because obesity was technically defined in terms of body fat per se, and BMI was derived only from height and weight.

In 2005, the Institute of Medicine (IOM) consciously departed from the terminology discussed above and elected to define children with a BMI at ≥95th percentile for age and gender as obese rather than overweight.<sup>44</sup> The IOM report expressed the seriousness, urgency, and medical nature of childhood obesity and deliberately sought to express this concern by using the term “obese” to refer to the children and adolescents with the highest BMI. A recent expert committee endorsed the IOM position and recommended to replace the terms “at risk of overweight” and “overweight” with the terms “overweight” and “obese,” respectively.<sup>4,29</sup> Accordingly, the expert committee recommended that individuals 2 to 18 years of age with a BMI of >30 kg/m<sup>2</sup> or ≥95th percentile for age and gender (whichever is smaller) should be considered obese. Individuals with a BMI at ≥85th percentile but <95th percentile or 30 kg/m<sup>2</sup> (whichever is smaller) should be considered overweight.

The expert committee believed that the terms “overweight” and “obese” better convey the seriousness and importance of the obesity epidemic to health providers, parents, and children and in a less ambiguous manner than the previous terms, although no specific literature was cited to support this view. Because BMI identifies the fattest individuals with acceptable accuracy, especially at the highest levels of BMI,<sup>45,46</sup> the expert committee believed that choosing more direct terms that may provide additional impetus for

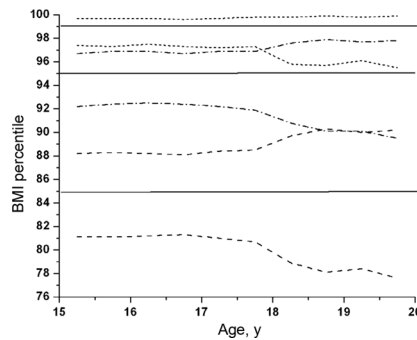
treatment and change was to be preferred to parsing technical concepts that would be unlikely to aid understanding. Finally, the new terminology comports with that from the IOTF BMI criteria for children and adolescents,<sup>47</sup> with conventional terminology for adults,<sup>32,48</sup> and with the *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM).

Nomenclature really does matter; it is a *sine qua non* with standardized definitions of health conditions. Standardized nomenclature increases precision in scientific and public communication and provides improved understanding in health guidance.

### Precision of Percentile Estimates

Often, health providers and researchers use the exact BMI-for-age cutoffs that define overweight and obesity as ironclad diagnostic criteria. Although standardized definitions are essential, as discussed above, the actual measurements on the child will always vary somewhat as a result of child and observer factors. In addition, the actual percentile cutoffs themselves are statistical estimates of points that are also subject to errors.

Let us assume that the basic data that were used to construct the 2000 CDC growth charts<sup>13</sup> are truly representative of the US population of children and adolescents, and that the statistical procedures used to smooth the percentile values across age were appropriate and unbiased. There still remains a degree of uncertainty regarding the point estimates of the final percentile values related to the number of children that were included in the samples within each 6-month age group used to estimate the percentiles. Simply put, the larger the sample, the more precise the percentile estimates, especially at the extremes of the distribution.



**FIGURE 1**  
The 85th, 95th, and 99th percentiles for BMI in girls (straight horizontal lines) and 95% CIs calculated from the number of subjects included at each age in the 2000 CDC growth charts.

As an example, Fig 1 presents the 85th, 95th, and 99th percentiles of BMI for girls as straight lines and the respective 95% confidence intervals (CIs) calculated by using the method of Wilcoxon<sup>49,50</sup> and the unweighted sample sizes within age groups (15–20 years) used for the 2000 CDC growth charts.<sup>13</sup> At younger ages the sample sizes range from 400 to 639, and the 95% CIs are quite stable and similar to those at 15 and 16 years. The sample sizes and corresponding 95% CIs for boys are similar to those for girls. The 99th percentile of BMI for age was not originally published with the growth charts but has been suggested as a useful cutoff for identifying children at added health risk.<sup>51</sup>

On the basis of sample sizes, the 95% CIs around the 85th BMI percentiles include values approximately between the 81st and 88th percentiles until ~17.5 years, when the sample sizes decrease and the 95% CIs become wider. These CIs mean that at 20 years of age (the most extreme case), a girl whose BMI percentile corresponds to the 85th percentile on the CDC chart may actually have a BMI anywhere between the 78th and 90th percentiles because of the imprecision of percentile estimates.

For the 95th BMI percentile estimates before ~17.5 years, the 95% CI range

from approximately the 92nd to the 97th percentiles but increase in spread, reaching from the 90th to 98th percentiles at the older ages. The 95% CIs around the 99th BMI percentiles for girls include from approximately the 97th to effectively just less than the 100th percentiles (because no point can exceed percentile 100).

After ~18 years of age, the upper 95% confidence limit for the 85th percentiles and the lower 95% confidence limit for the 95th percentiles are approximately coincident, and the upper limit of the 95th and the lower limit of the 99th percentiles actually overlap. This means, for example, that a 19-year-old girl with a BMI identified as being at the 99th percentile by the 2000 CDC growth charts (or by computer programs that calculate the exact percentiles) will probably have a BMI somewhere between the 96th and 100th percentiles.

The precision of the upper percentile cutoffs for BMI can be viewed from several different perspectives. First, the samples and CDC growth charts are as they are, and no revisions are anticipated in the near future. Consequently, those who use the growth charts should understand their limitations in interpreting findings and not wait for more precise estimates. Actually, additional imprecision beyond that related to sample size probably also occurs at some ages in adolescence because of differences in maturational status.<sup>52</sup>

Given the range of CIs surrounding the BMI percentiles at all ages, health providers and investigators should be a little less stringent in defining the exact location of a child or groups of children in the BMI distribution relative to the growth charts. Accordingly, BMI values just below or just above recommended cutoffs should be interpreted as only 1 indicator and not the only diagnostic criterion for clinical decisions. Follow-up visits and repeated as-

assessments on other occasions should reduce the uncertainty of the child's BMI status.

The fairly wide confidence limits around the percentiles do not invalidate the recommended BMI cutoffs for standardized reporting of population prevalences or for analyses of the associated risk profiles of groups of children.<sup>51</sup> Nevertheless, investigators should be cautious drawing inferences from risk ratios comparing the observed and expected prevalences beyond a given BMI cutoff because of the imprecision of the percentile cutoffs.

### When Should *z* (SD) Scores Be Used?

A BMI *z* or SD score is the BMI of a child transformed into a scale comprising the number of SD units it is away from the mean of the referent population of the same age and gender. The 2000 CDC growth charts<sup>13</sup> were constructed in such a way to allow calculation of *z* scores for BMI.

There are several advantages to using *z* scores compared with using the corresponding percentiles, although they both describe a child's status relative to the same reference data set. The pediatric applications of *z* scores that are most common are probably in endocrinology or nutrition where children who are very small relative to the growth charts are seen and *z* scores provide a more useful and manageable metric than percentiles to evaluate and monitor status or treatment.<sup>53,54</sup> For example, a 3-year-old boy with a height-for-age *z* score of  $-3.6$  has a height that is 3.6 SDs lower than the age- and gender-specific mean for him on the growth charts; his corresponding height-for-age percentile is 0.013.

When a high proportion of children have heights and weights less than the lowest percentiles (eg, 3rd, 5th), as found in many developing countries,

the percentile charts cease to be useful for differentiating their growth status. Accordingly, cutoffs of less than  $-2$  *z* for height for age and weight for age have become conventional definitions for stunting and wasting, respectively.<sup>55</sup>

In a similar fashion, for overweight and obesity in children and adolescents, *z* scores can be useful for characterizing individuals with a high BMI that exceeds the percentile levels available on the growth charts. For example, if the progress of a girl with a BMI that far exceeds the 97th percentile for age (currently the highest percentile available on the CDC charts) is monitored, her attained BMI on the growth chart is difficult to evaluate and impossible to meaningfully quantify. On the other hand, by converting her BMI to a *z* score, her progress can be monitored and changes in subsequent *z* scores have a direct interpretation relative to the referent population of her age. Because *z* scores are calculated relative to age, noting a change in *z* score is an appropriate way to evaluate changes in BMI across ages relative to what is expected in the referent population.

An alternative to using *z* scores to evaluate change in individual children with elevated BMI is to just use change in BMI itself. These changes are understandable to practitioners, adolescents, and families, and they allow setting of goals and monitoring of progress.

Using *z* scores is currently the only appropriate way available to quantify the severity of obesity in children who have BMI levels that exceed the available percentiles for age and gender. Unfortunately, *z* scores require a computer program to calculate them readily, and the SD-related metric is not familiar to many practitioners. Because the total variation in BMI (eg, the distance between the 5th and 95th percentiles) progressively increases with

age, calculating the percentage excess of a BMI value or percentage overweight beyond a percentile value is inappropriate, because it will have inconsistent meaning from age to age.

### CHALLENGES OF MEASURING HEIGHT, WEIGHT, AND BMI

Summary reminders concerning data collection and management are listed in Table 1. The particular setting in which data for BMI assessments will be collected has implications for how or whether the recommended practices can be implemented.

#### Equipment and Space

If possible, height should be measured to the nearest 0.1 cm ( $\frac{1}{4}$  in) by using a stadiometer mounted on the wall or a portable stadiometer that allows the child to be positioned properly with his or her back against a vertical surface. A second choice are models that measure the child freely standing, but the measurement errors for these latter instruments tend to be larger than when the measurements are taken with the child standing against a surface.<sup>55</sup> The height measurements for the 2000 CDC growth charts<sup>13</sup> were taken by using wall-mounted stadiometers. Many brands of acceptable stadiometers are available, and searching on-line will provide several good choices. Stadiometers attached to scales that do not allow the child to be positioned correctly are not recommended.

Weight should be measured by using a good-quality scale to the nearest 100 g ( $\frac{1}{4}$  lb). In the past, balance-beam scales were routinely recommended because the only alternatives were spring scales that were less dependable. Now there are many good electric scales available that are also quite portable. The more expensive scales have multiple pressure transducers under the weighing platform, so they are less sensitive



**TABLE 1** Data-Collection and Management Practices for Reducing Errors for Height, Weight, and BMI

Equipment and space
Choose appropriate equipment
Check and calibrate equipment regularly
Keep extra batteries for scales
Provide a private area for child measurements, if possible
Measurement protocols
Choose a protocol that matches that used in the growth charts
Have written copies of measurement protocols available for review
Train and standardize data collectors
Make sure data are recorded in the appropriate units (eg, kilograms, pounds)
Make sure data are measured and recorded to the nearest unit specified in the protocol (eg, 0.1 cm for height, 0.1 kg for weight)
Collect some replicate measurements for assessment of reliability, if feasible
Personnel
Use as few observers as is feasible to take measurements, especially for research studies
Identify observers on data-collection forms or data-entry programs
Data management
Use as exact ages as possible
Have unique identifiers for children
Calculate BMI, percentiles, and z scores by using tables or computer programs

to variation in the child's position and shifting of weight from one leg to the other. Again, an Internet-based search will yield many good alternatives.

In a research setting, obviously, the best-quality equipment should be chosen for maximum consistency over time and for reliability among observers taking the measurements. In clinical or community settings, cheaper alternatives are often used, but given the heavy utilization in a busy clinic, for example, investing in sturdy anthropometric equipment that can be calibrated if necessary will prove worthwhile and increase confidence in the measurements. Cheaper models of stadiometers tend to have less-rigid parts that wobble or bend with frequent use.

With repeated use or if equipment is moved about fairly often, stadiometers and scales should be checked to determine if they are calibrated correctly. It is important to develop a regular schedule for calibration (eg, daily in research, weekly in clinic) and assign someone to be responsible for these duties. Depending on the installation, good stadiometers usually can be cal-

ibrated by using a metal rod of a fixed length.

Good electric scales can be calibrated or "zeroed." In most areas of the United States, state agencies in departments of commerce, standards, or agriculture have representatives who calibrate and certify scales in grocery stores and in other commercial venues. In some cases, these representatives can be called on to routinely check and calibrate scales at permanent sites. Alternatively, scales can be calibrated by using weights of known size. If models of electric scales are used in clinic or in the field, ensuring that a supply of batteries of appropriate size should be on the checklist for routine equipment maintenance.

Often, in busy clinic or school situations, stadiometers and scales are relegated to hallways or even reception areas. Children and adolescents may find it embarrassing to be measured, and even more so to have witnesses to the procedures.<sup>56</sup> Having a private or partially screened area for the height and weight measurements will increase child cooperation and enhance

the patient confidentiality sought by institutional human subjects committees.

### Measurement Protocols

Because health providers and others who use BMI data will almost always compare them to the growth charts, it makes sense to strive to collect the height and weight measurements that comprise BMI by using protocols that match those used in the reference data as closely as possible. The measurement procedures used in the collection of the height, weight, and BMI data for the 2000 CDC growth charts<sup>13</sup> are currently available as a downloadable file at the CDC National Health and Nutritional Examination Survey (NHANES) Web site ([www.cdc.gov/nchs/data/nhanes/bm.pdf](http://www.cdc.gov/nchs/data/nhanes/bm.pdf)). These measurement protocols follow closely those recommended by a US consensus group.<sup>57</sup> This publication has become the gold-standard reference in the United States for anthropometry methods related to health issues, although slight differences exist for some measurements customarily used internationally.<sup>58</sup>

It is important to train data collectors in the appropriate methods for measuring height and weight. Again, the goal is to use the same measurement protocols that were used for the derivation of the growth charts. Sometimes, experienced clinic staff may take offense because they have been measuring height and weight for a long time. Often, however, "the way we do it here" includes some bad habits or deviations from the prescribed protocols. Standardizing all data collectors to a gold-standard trainer ensures that a single protocol is followed and that departures from the trainer are within acceptable limits.<sup>59</sup>

For extended research protocols or for ongoing surveillance or clinical activities, having a gold-standard trainer periodically visit and observe measurements or take some replicate

measurements will help prevent “drift” in the measurement techniques. Also, these opportunities can be used to correct and recertify data collectors, if necessary.

Laminated copies of the measurement protocols on-site provide a readily available reminder for data collectors concerning child position, measurement landmarks, and local policies regarding calibration, clothing, exclusion criteria, data recording, data flow, etc.

In research settings a certain proportion of the measurements should be repeated to evaluate measurement reliability. The proportion required depends on the number of different observers concerned, the number of children usually measured, and the period over which reliability will be assessed. In general, there should be enough replicates to capture the variation among data collectors and study design features and to capture a fairly stable estimate of the mean differences between replicates and the accompanying SD. The SD of differences between replicates is really a measure of variance, and the CIs for a variance begin to stabilize at sample sizes larger than 20 (in our case, 20 pairs of measurements).<sup>60</sup>

As an example, for a hypothetical study in school-aged children, a 3-person measurement team will visit 4 different schools during a month of data collection. Each school has an average of 30 children, and the team will average ~10 children measured per day. So, each school will require 3 days of data collection, and ~120 children will be measured. If a target of 25 replicates is sought for assessing measurement reliability, that amounts to an ~20% sample. One simple approach is to specify that the data collectors remeasure 2 children per day and that a different data collector from the one who measured the child the

first time take the measurements. Over the course of the month of data collection, the variation among observers, schools, and any study drift will be captured in the final reliability sample, which should include data on ~25 children. For complicated protocols that involve many measurements or administration of other instruments, children may contribute only a replicate for one of the measurements so that the burden on any one child is small and the total of 25 replicates may represent many more individual children. The calculation of the relevant measurement-reliability statistics has been explained elsewhere.<sup>27,61</sup>

If the measurement protocols specify that duplicate measurements be routinely collected for all subjects (as recommended above), then these replicates can be used for assessing measurement reliability as long as all the different data collectors involved in the study take the replicates. If the mean of replicate measurements will be used in statistical analyses, the measurement reliability should take this into account.<sup>61</sup> If different data collectors usually work on different days, then special scheduling may be required to accommodate fully capturing the interobserver variation in the reliability sample.

### Personnel

Experience shows that an advanced formal education is not required to take high-quality anthropometric measurements. Willing adults who will give adequate attention to detail and who meet the requirements for employment are usually satisfactory. Members of the community who are familiar with the local ethos and jargon may be excellent data collectors. In some situations, like-gender observers may make children and adolescents more comfortable with the

touching required for anthropometric measurements.

As mentioned previously, having as few data collectors as is feasible for other practical demands will minimize interobserver measurement variation. Ensuring that unique observer codes are included on the data-collection forms or data-entry computer programs will aid in quality-assurance activities and can even be used in the statistical analyses if consistent observer measurement bias becomes apparent.

### Data Management

Having chronological ages as exact as possible is important for the accurate calculation of percentiles and z scores, and they will aid in minimizing age-related variance in statistical analyses when children are grouped according to age. Chronological ages in years expressed to at least 2 decimal points are sufficient for most applications; this will capture exact ages to the nearest 3 days. For children less than 5 or 6 years of age it may be more convenient to express age in months to 1 decimal point, or in exact days.

Actual values for BMI, BMI percentiles, and BMI z scores are best calculated by using computer programs to avoid computational errors. There are many Web sites with BMI calculators that can be found easily by using Internet searches, including those provided by the CDC (<http://apps.nccd.cdc.gov/dnpabmi/Calculator.aspx>) and National Institutes of Health ([www.nhlbisupport.com/bmi/bminojs.htm](http://www.nhlbisupport.com/bmi/bminojs.htm)).

In some settings where immediate patient feedback or charting are conducted, calculating BMI by using tables may be preferred. Again, many Web sites provide such tables; the only caution is that some BMI tables are designed for adults and may not include the low heights and weights observed in children.<sup>2</sup>

**TABLE 2** Selected Studies With Interclass Correlation Coefficients for Reported and Measured Height, Weight, and BMI According to Gender

Source	Group/Location	Age or Grade Level	n, Male/Female	Height <i>r</i>		Weight <i>r</i>		BMI <i>r</i>	
				Male	Female	Male	Female	Male	Female
Davis and Gergen <sup>63</sup>	Mexican American/US	12–19 y	392/437	0.86	0.86	0.95	0.93	0.87	0.85
Himes and Faricy <sup>64</sup>	All/US	12–16 y	759/876	0.89	0.79	0.97	0.93	0.93	0.87
Himes and Story <sup>65</sup>	American Indian/Minnesota	12–19 y	41/28	0.91	0.71	0.96	0.91	0.90	0.80
Hauck et al <sup>66</sup>	American Indian/US	12–19 y	536	0.83	0.62	0.95	0.90	0.88	0.79
Brener et al <sup>67</sup>	20 states/US	Grades 9–12	957/1075	0.87	0.82	0.92	0.94	0.89	0.89
Himes et al <sup>68</sup>	Minnesota/US	12–18 y	1936/1861	0.90	0.80	0.96	0.94	0.89	0.85
Tsigilis <sup>69</sup>	Trikala, Greece	Middle and high schools	141/159	0.94	0.93	0.97	0.97	0.90	0.94
Median <i>r</i>				0.89	0.80	0.96	0.93	0.89	0.85

Exact BMI percentiles and BMI z scores can be calculated by using Epi Info, a free, user-friendly and downloadable computer program developed by the CDC ([www.cdc.gov/epiinfo](http://www.cdc.gov/epiinfo)). At the CDC Web site, researchers can download a program for SAS statistical analysis software that generates a data set containing the percentiles and z scores for all the anthropometric measurements (including BMI) in the 2000 CDC growth charts ([www.cdc.gov/nccdphp/dnpa/growthcharts/resources/sas.htm](http://www.cdc.gov/nccdphp/dnpa/growthcharts/resources/sas.htm)).

**SELF-REPORTED HEIGHT, WEIGHT, AND BMI**

Having older children and adolescents report their height and weight rather than having someone directly measure them is attractive economically and logistically. Costs of direct anthropometric measurements include additional time, personnel, training, and equipment. Logistically, direct measurements require an in-person examination, space, and additional time for participants. If direct measurements of height and weight are required, some study designs and data-collection strategies are summarily inadequate or eliminated (eg, mail surveys, classroom surveys, telephone surveys). Of course, the appropriateness of using self-reports of height, weight, and BMI depends on the reliability, bias, validity, and specific applications of these measures. In some cases, self-reported data may be all that exist, so it is important to understand

when and how such data might be used appropriately.<sup>62</sup>

**Measurement Reliability**

No published data are available on reliability in self-reported height and weight as narrowly defined previously (ie, the random error associated with the same measurement being repeated). Such data would comprise the same children being asked for their reported height and weight at least twice over a period of time insignificant for growth.

Reliability in self-reports has been evaluated in adolescents, considering reliability as the random errors associated with the differences between self-reported height, weight, and BMI and the corresponding measured dimensions. A good summary measure of this reliability is the Pearson or interclass correlation coefficient.

Correlation coefficients between reported and measured height, weight, and BMI are presented in Table 2 for some selected studies that reported the correlations according to gender. Overall, the correlations for reported and measured dimensions are relatively high, indicating that self-reported values are generally reasonable proxies for the corresponding measured values. On the basis of the correlation coefficients, boys generally do a little better than girls, and weight is usually more reliably reported than is height. Because self-reported BMI combines

the random errors in both height and weight, self-reported BMI generally has lower correlations with measured BMI than corresponding associations observed for reported and measured height and weight.

The youngest-aged children included in these studies were 11 to 12 years old, and correlations between self-reported and measured dimensions, especially height, are usually lower at these ages than they are later in adolescence.<sup>64,70,71</sup>

A slightly different concern about young adolescents is that they are often unable or decline to report their heights and weights.<sup>62,72</sup> In a study based on US national-level data, 41% of 12-year-olds and 25% of 13-year-olds had missing data for weight.<sup>64</sup> These rates compared with 4% missing reported weights in 15- and 16-year olds. It may be that for youth aged 11 to 13 years their height has not yet become as important to them as it will be as they get older, and they may not have regular opportunities to have their height measured.

**Measurement Bias**

Although Pearson correlation coefficients are useful indicators of reliability, they only provide average associations, and they only account for random errors between reported and measured values. Pearson correlations are blind to systematic errors or bias. Several different sources of bias

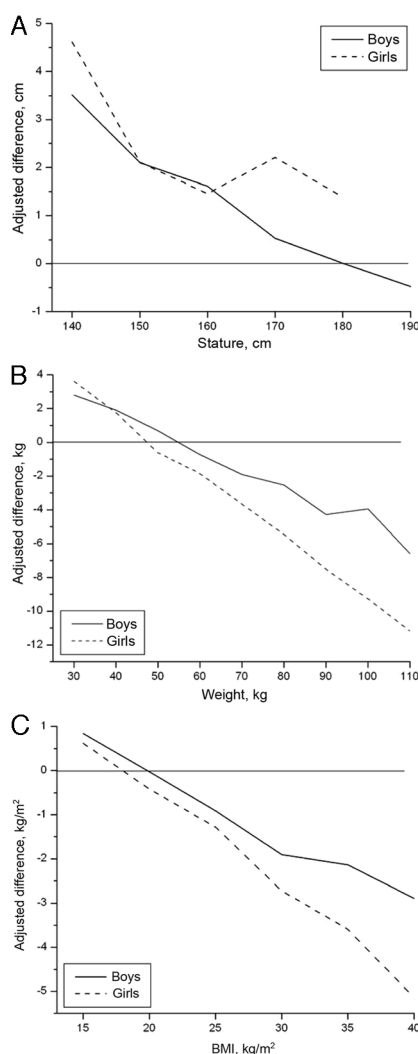
in self-reports of height and weight have been investigated, and they were recently reviewed for studies on US adolescents.<sup>62</sup>

For our discussion, it is important to recognize that the mean values of self-reported height are usually overestimated by ~1 to 2 cm, and mean self-reported weight is usually underestimated by 2 to 4 kg, especially so in girls.<sup>62,72</sup> Thus, with overestimated height and underestimated weight, mean BMI values calculated from the self-reported data are usually less by 2 to 3 BMI units (kg/m<sup>2</sup>) than if they were measured.

Another source of bias that is important for understanding how self-reported data might be used in evaluating overweight and obesity is related to the body size of the children and adolescents providing the self-reports. The mean differences for self-reported values less measured values for height, weight, and BMI are presented in Fig 2 relative to categories of the measured dimensions for a sample of 3797 Minnesota youth aged 12 to 18 years.<sup>68</sup>

For height, the errors in self-reporting are largely positive because most of the youth overestimated their heights (mean differences: boys, 1.2 cm; girls, 2.4 cm). Nevertheless, a strong negative relationship between the errors in reporting height and the actual measured heights is evident so that the only group actually underestimating height was the very tallest boys. For self-reported weight and BMI, the errors in self-reports became increasingly negative (indicating underestimates) as categories of measured weight and BMI increased, with steeper slopes in girls than in boys.

This pattern of underestimation means that the greatest impact of the bias in self-reported BMI will be to underestimate prevalences of overweight and obesity defined by the



**FIGURE 2** Mean differences between self-reported and measured body size adjusted for age, socioeconomic status, and race/ethnicity, plotted against categories of the measured dimension: A, height; B, weight; C, BMI.<sup>68</sup>

upper percentiles (eg, 85th, 95th). For example, in a separate study of high school students by Brener et al,<sup>67</sup> the prevalences for overweight ( $\geq 85$ th percentile) were 47.4% for directly measured BMI and 29.7% for self-reported BMI. Corresponding prevalences for obesity ( $\geq 95$ th percentile) were 26.0% for measured BMI and 14.9% for self-reported BMI. Unfortunately, there is no easy conversion from a prevalence based on self-reported BMI to what it would have been if height and weight were measured.

The biases in self-reports are entangled in idiosyncratic differences among samples in gender, age, underlying distributions of BMI, and perhaps race, mental health, and socioeconomic status.<sup>62,68</sup>

### Measurement Validity

From the evidence for bias discussed above, it is not surprising that considerable misclassification occurs when children and adolescents are identified as overweight or obese on the basis of self-reports and the BMI-percentile criteria. In the Brener et al<sup>67</sup> study, the sensitivity and specificity of self-reported BMI for identifying overweight adolescents were 60.5% and 98.0%, respectively. Corresponding values for sensitivity and specificity for identifying obese individuals were 54.9% and 99.2%, respectively. So, as few as 55% (positive predictive value) of those who are truly overweight will be correctly identified as such when using BMI calculated from self-reported heights and weights. Results from other studies of validity are not much more encouraging.<sup>62</sup>

The validity of BMI using self-reported data relative to total body fat has not been evaluated. Nevertheless, given the modest validity relative to measured BMI, BMI derived from self-reported data must be even poorer than measured BMI in its ability to correctly identify the fattest individuals on the basis of laboratory methods.

### When Is It Appropriate to Use Self-reported BMI?

In some situations, BMI derived from self-reported data are the only data available (eg, the CDC Youth Risk Behavior Surveillance System,<sup>73</sup> which collects data through telephone interviews from a national sample). In other cases, the complexity and size of the survey make direct measurements impractical.<sup>74</sup> Nevertheless, any use of



self-reported height, weight, and BMI should be done with an understanding of their limitations and biases despite the obvious logistic and economic advantages. Interpretation of findings needs to be couched accordingly.

For surveillance purposes, prevalences of child and adolescent overweight and obesity are important for describing the nature and extent of problems, monitoring trends or changes over time, and comparing communities or regions for program priorities. An important notion to understand here is that prevalences of overweight and obesity based on self-reported BMI data will almost certainly be underestimates of the true prevalences, although to an unknown degree. Consequently, prevalences based on self-reported data should not be the basis of determining the extent of local problems compared with prevalences reported in the scientific literature describing national patterns and trends and based on measured height and weight.

Although there are no specific data to elucidate whether biases in self-reports are fairly stable over time, it is not unreasonable to assume that factors related to reporting biases should not dramatically change within the same group over relatively short time periods (eg, 1–2 years). Consequently, with that assumption, it should be acceptable to use prevalences from self-reported data to assess changes over time within groups. When using the same rationale, self-reported data may be acceptable to use in some program evaluations, provided that interventions do not include behavioral or psychological components that may alter body awareness or self-image that may be related to the biases in child reporting.<sup>75</sup>

Comparing prevalences of overweight or obesity among different locations or groups by using self-reported data

are problematic, because the comparison assumes that all the possible factors that contribute to the biases in reporting are the same, including the underlying distribution of measured BMI. Hence, if one concludes that there are meaningful differences in prevalences of child obesity on the basis of self-reports, one had to arrive at that conclusion by assuming that the underlying distribution of BMI was the same.

Self-reported height, weight, and BMI should not be used to assess body size in clinical settings where diagnostic and therapeutic decisions are made. The individual variation in self-reported values is impossible to predict, and the consequences of misclassification may be serious. Certainly, for most research protocols, directly measured height, weight, and BMI are strongly recommended because they increase the precision and accuracy of the estimates and they avoid the need to make assumptions regarding BMI status due to unmeasured factors.

### PARENT-REPORTED HEIGHT, WEIGHT, AND BMI

If parents were able to accurately report the height and weight of their young children, it would have many of the economic and logistic advantages proposed for self-reports in older children. Unfortunately, there are far fewer data directly evaluating the validity of parental reports of child height and weight compared with those available for self-reports, and the results are more difficult to generalize because of differing study protocols and analyses, disparate child ages, and sometimes conflicting or even confusing results.

For parents to have reasonably accurate knowledge of their child's height and weight usually requires that they have either measured the child themselves or been informed from the school or clinic where someone else

measured them. Obviously, the more recently the measurements were taken, the more accurate the parent reports should be.

The fact that children continue to grow after the most recent measurement at home or clinic has led some to conclude that parental reports will always underestimate the measured height and weight of children. Nevertheless, for most studies that make direct comparisons, the mean parent-reported child height and weight were close to the corresponding measured means, usually within  $\pm 1$  cm or  $\pm 1$  kg,<sup>76–80</sup> and reasonably represented by underestimates and overestimates of the measured means. There are, however, exceptions. Mexican American mothers who participated in the US Hispanic Health and Nutrition Examination Survey (HHANES) consistently underestimated the mean measured height of their children (6 months to 11 years) by 6 to 9 cm while reporting mean weight within 1 kg of the mean measured weight.<sup>65</sup> Almost one fourth of the Mexican American mothers said that they did not know their child's height and weight and, thus, were unable to report it. For a sample of 818 Spanish children aged 6 to 8 years, mean parental-reported child height was 2.4 cm taller than the mean measured height,<sup>81</sup> and mean measured weight was slightly overestimated but, again, within 1 kg.

The net effect on mean BMI calculated from parental reports compared with the mean of measured BMI in the available studies is accordingly usually small and within  $\pm 1$  kg/m<sup>2</sup>, except for the Mexican American mothers in the US Hispanic Health and Nutrition Examination Survey, who substantially underestimated child height in their reports and, therefore, overestimated the mean BMI calculated from measured height and weight.<sup>65</sup>

Although the average biases in parental reports of height, weight, and BMI tend to be rather small, there could still be systematic biases in parental reports according to the measured size of the children, as occurs with self-reports in older children. The few studies that have investigated this question indicated that parental reports of child weight tend to overestimate the lightest children and underestimate the heaviest children, or a regression toward the overall measured mean.<sup>63,76,80,82</sup>

From these reporting biases related to measured child size, one should expect the prevalences of child overweight and obesity based on parental reports to systematically underestimate corresponding prevalences based on measured BMI. Nevertheless, the studies that have found both prevalences of obesity are about equally represented by those that found parental reports to yield higher prevalences of obesity compared with estimates based on measured BMI<sup>81,82</sup> and those that found parental reports to yield lower prevalences of obesity.<sup>79,80</sup> In a study that included parents of Japanese children in the first and fourth grades, the differences between prevalences of obesity based on parental reports and measured BMI were small but in both directions when reported separately according to gender.<sup>77</sup>

From the available literature one must conclude that prevalences of child obesity based on parental reports do not differ systematically or dramatically from corresponding prevalences based on direct measurements. In all of these studies differences between prevalences of childhood obesity determined by measurements and parent reports were small, usually within  $\pm 5\%$  (in the absolute prevalence).

After reviewing the available literature on parental reports of height and weight,

a few summary recommendations emerge relative to appropriateness for studies and practices related to child overweight and obesity. First, of course, is that BMI derived from directly measured height and weight is always preferable to similar data obtained from parental reports. Certainly, in clinical settings where decisions concerning diagnosis or treatment are made, BMI assessments should be made by using directly measured height and weight. The probability of misclassification of individuals relative to overweight or obesity status is simply too great to use parental reports of child height and weight.

As when choosing to use BMI data calculated from self-reports of height and weight for any purpose, BMI estimates from parental reports always should be used with less certainty than corresponding data obtained from direct measurements. One dimension of the uncertainty is that the direction of the biases using parental reports compared with BMI estimates from direct measurements are poorly understood, so one cannot reasonably conclude that the true mean BMI or prevalence of obesity is more or less than that obtained from parental reports.

Accordingly, important decisions regarding obesity trends and programs for groups should not be based only on comparisons of prevalences of child overweight and obesity based on BMI derived from parental reports with those from national or regional surveys that used direct measurements of height and weight for estimates of BMI. While acknowledging the uncertainty of BMI based on parental reports, it does not seem unreasonable to compare prevalences of overweight or obesity when similar approaches for obtaining the parental reports have been used for the estimates when such data are all that are available.

A strong recommendation is that those who use parental reports routinely in their surveys develop research to validate this approach and to document its reliability and validity. The available literature is simply too sparse to allow any reasonable conclusions to be drawn regarding specific factors that may be related to unreliability or biases in parental reports (eg, child age, gender of the parent, prevalence of obesity, socioeconomic status, BMI of parents, etc).

### SKINFOLD THICKNESS

A skinfold thickness is the double layer of skin and subcutaneous fat (panniculus adiposus) lifted as a fold and measured with standardized calipers and methodology at specific sites on the body.<sup>83</sup> The rationale for measuring skinfolds as an indicator of overweight and obesity is that subcutaneous fat is part of and highly correlated with total body fat. Skinfolds have a long history of use as an indicator of nutritional status and body fatness, and their validity and measurement properties are well established.<sup>84,85</sup>

Skinfold-thickness measurements generally are more highly correlated with total body fatness than is BMI,<sup>28,51,86</sup> although the association varies by the degree of body fatness. Skinfold-thickness cutoffs can correctly identify the fattest children about as well as BMI.<sup>46,87,88</sup> Nevertheless, for the present discussion, skinfolds are of interest if their addition improves the accurate identification of children as overweight or obese beyond that provided by BMI alone. There are few data addressing this specific question, although Mei et al<sup>46</sup> found that triceps or subscapular skinfold measurements greater than a series of cutoffs failed to improve the identification of the fattest children (by dual-energy radiograph absorptiometry) beyond that achieved by using a 95th-percentile

cutoff for BMI. So, children with a BMI at the tail of the population distribution (ie,  $\geq 95$ th percentile) have little misclassification as the fattest because there is little “hidden” muscularity or fatness.

There are several practical reasons that make skinfolds challenging to use. Measurement reliabilities for skinfold measurements are usually much lower than for height and weight,<sup>27</sup> even in well-trained hands. Attaining the maximum measurement reliabilities for skinfold measurements requires substantial experience and regular practice, probably more than most personnel have in clinical and community settings. Finally, there are no published reference percentiles available for skinfold thickness for US children, so no optimum percentile cutoffs have been defined.

In sum, skinfold-thickness measurements remain important in many research applications, but they cannot be recommended as a routine part of screening, management, or surveillance of child and adolescent overweight and obesity.

## WAIST CIRCUMFERENCE

The measurement of waist circumference is an attempt to capture information regarding the distribution of body fat, in this case the visceral adipose tissue that has been linked to increased health risks and metabolic disorders in children and adults.<sup>89,90</sup> In multiple regression models, waist circumference (as a continuous measure) does better than BMI in predicting insulin resistance, blood pressure, serum cholesterol levels, and triglyceride levels,<sup>91–93</sup> especially in adolescents. Also, the ratio of waist circumference to height has been shown to be associated with cardiovascular risk factors.<sup>94</sup>

Nevertheless, when it comes to accurately identifying the fattest children,

thresholds of waist circumference do no better than those of BMI or triceps skinfold thickness.<sup>87</sup> There is no information regarding any additional benefit of using waist circumference to identify the fattest children once BMI criteria have been applied.

Waist circumference is easier to measure reliably than skinfolds, and interobserver measurement reliabilities are usually intermediate between those for height and weight and those for skinfold thicknesses.<sup>27</sup> Age-based percentiles of waist circumference are available for US children,<sup>95</sup> but no optimum cutoffs for identifying the fattest children have been developed.

The main utility of waist circumference per se is as a measure of fat distribution rather than total body fatness. Consequently, although some of the findings regarding identifying adolescents at risk for concurrent or future morbidity are important, there is little to indicate that including waist circumference adds appreciably in the identification of the fattest individuals beyond what is available with BMI. Accordingly, waist circumference is not recommended to be routinely included in screening and surveillance for child and adolescent overweight and obesity.

## PERCENTAGE BODY FAT FROM BIOELECTRICAL IMPEDANCE

Bioelectrical impedance analysis (BIA) measures the opposition of body tissues to a small ( $<1$ -mA) alternating current that is imperceptible to the subject. Because bioelectrical impedance differs between lean tissue (because of its water content) and fat tissue, BIA combined with body height may be used to estimate body water, fat-free mass, and body fat.<sup>96</sup> Total body fat may then be divided by body weight and multiplied by 100 to yield the percentage of body weight that is fat. BIA has gained popularity because

it is noninvasive, portable, and reliably measured.<sup>97</sup> Measurement reliabilities for BIA are generally high and can approach those for height and weight.<sup>98</sup>

The best studies of the validity of BIA in estimating total body fat use empirical equations that relate the measured resistance and reactance and child anthropometry to an independent criterion measure of total body fat. It turns out that the BIA prediction equations are closely tied to the referent population, so different equations can yield different body-fat estimates for children in other samples with the same BMI.<sup>99</sup> Consequently, although measured with good reliability, biases resulting from prediction equations that just do not fit right in the population studied are a concern with BIA.

With the introduction of simple foot-to-foot BIA assessments that just require that the child step on scales with electrode foot plates, BIA has become increasingly popular because the child is not required to lay quietly supine for the procedure. Nevertheless, some models of the foot-to-foot BIA apparatus only provide the summary measure of percentage body fat, and the details of the resistance, reactance, equations, or the referent population are not available. Accordingly, one is left to only hope that the equations are appropriate in a particular population. In a study of overweight and obese adolescents, foot-to-foot estimates underestimated total percent body fat by 2% to 3% body fat.<sup>100</sup>

Another practical concern for using BIA to estimate overweight and obesity is that there are no percentile reference data for percent body fat in US children and no criterion thresholds established to identify those at greatest health risk. Means and SDs for percent body fat for US adolescents older than 12 years derived from BIA have been published.<sup>101</sup>

With BIA, then, one is left with a technically excellent method but with practical constraints (need for appropriate equations and reference data) that limit its usefulness in most settings where identification and management of overweight and obesity are conducted. Consequently, BIA is not recommended for routine use in addition to assessment using BMI. More detailed descriptions of BIA and other measures of body composition have been elaborated elsewhere.<sup>97,102</sup>

## SUMMARY AND CONCLUSIONS

BMI is an important indicator of overweight and obesity in childhood and adolescence. When measurements are taken carefully and compared with

appropriate growth charts and recommended cutoffs, BMI provides an excellent indicator of overweight and obesity sufficient for most clinical, screening, and surveillance purposes.

Accurate measurements of height and weight require that adequate attention be given to data collection and management. Choosing appropriate equipment and measurement protocols and providing regular training and standardization of data collectors are critical aspects that apply to all settings in which BMI will be measured and used.

Proxy measures for directly measured BMI, such as self-reports or parental reports of height and weight, are much less preferred and should only be used with caution and cognizance of the

limitations, biases, and uncertainties attending these measures.

There is little evidence that other measures of body fat such as skinfolds, waist circumference, or bioelectrical impedance are sufficiently practicable or provide appreciable added information to be used in the identification of children and adolescents who are overweight or obese. Consequently, for most clinical, school, or community settings these measures are not recommended for routine practice. These alternative measures of fatness remain important for research and perhaps in some specialized screening situations that include a specific focus on risk factors for cardiovascular or diabetic disease.

## REFERENCES

1. Freedman DS, Sherry B. The validity of BMI as an index of body fatness and risk among children. *Pediatrics*. 2009;124(3 suppl):S23–S34
2. Himes JH, Dietz WH. Guidelines for overweight in adolescent preventive services: recommendations from an expert committee. The Expert Committee on Clinical Guidelines for Overweight in Adolescent Preventive Services. *Am J Clin Nutr*. 1994;59(2):307–316
3. 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). Available at: [www.pediatrics.org/cgi/content/full/102/3/e29](http://www.pediatrics.org/cgi/content/full/102/3/e29)
4. Krebs NF, Himes JH, Jacobson D, Nicklas TA, Guilday P, Styne D. Assessment of child and adolescent overweight and obesity. *Pediatrics*. 2007;120(suppl 4):S193–S228
5. Nunnally JC. *Psychometric Theory*. 2nd ed. New York, NY: McGraw-Hill Publishing Company; 1978
6. Zeller RA, Carmines EG. *Measurement in the Social Sciences: The Link Between Theory and Data*. Cambridge, United Kingdom: Cambridge University Press; 1980
7. Rao DH, Sastry JG. Day-to-day variation in body weights of children. *Ann Hum Biol*. 1976;3(1):75–76
8. Lampl M, Birch L, Picciano MF, Johnson ML, Frongillo EA Jr. Child factor in measurement dependency. *Am J Hum Biol*. 2001;13(4):548–557
9. Wasse M. Part of a letter from the Reverend Mr. Wasse, Rector of Aynho in Northamptonshire, to Dr. Mead, concerning the difference in the height of a human body between morning and night. *Philos Trans (1683–1775)*. 1724;33:87–88
10. Roberts N, Hogg D, Whitehouse GH, Dangerfield P. Quantitative analysis of diurnal variation in volume and water content of lumbar intervertebral discs. *Clin Anat*. 1998;11(1):1–8
11. Strickland AL, Shearin RB. Diurnal height variation in children. *J Pediatr*. 1972;80(6):1023–1025
12. Voss LD, Bailey BJ. Diurnal variation in stature: is stretching the answer? *Arch Dis Child*. 1997;77(4):319–322
13. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, et al. CDC growth charts: United States. *Adv Data*. 2000;(314):1–27
14. Khosla T, Billewicz WZ. Measurement of change in body-weight. *Br J Nutr*. 1964;18(PubMed did not provide):227–239
15. Marshall EL. A review of American research on seasonal variation in stature and body weight. *J Pediatr*. 1937;10:819–831
16. Brown KH, Black RE, Becker S. Seasonal changes in nutritional status and the prevalence of



- malnutrition in a longitudinal study of young children in rural Bangladesh. *Am J Clin Nutr.* 1982;36(2):303–313
17. Maleta K, Virtanen SM, Espo M, Kulmala T, Ashorn P. Seasonality of growth and the relationship between weight and height gain in children under three years of age in rural Malawi. *Acta Paediatr.* 2003;92(4):491–497
  18. Bransby ER. The seasonal growth of children. *Med Off.* 1945;73:149–151; 157–159; 165–168
  19. Cole TJ. Seasonal effects on physical growth and development. In: Ulijaszek SJ, Strickland SS, eds. *Seasonality and Human Ecology: 35th Symposium Volume of the Society for the Study of Human Biology.* Cambridge, United Kingdom: Cambridge University Press; 1993:89–106
  20. Ikeda M, Watanabe T. Constant growth of primary school children throughout four seasons of year. *Tohoku J Exp Med.* 1985;145(4):413–418
  21. Roche AF, Himes JH. Incremental growth charts. *Am J Clin Nutr.* 1980;33(9):2041–2052
  22. von Hippel PT, Powell B, Downey DB, Rowland NJ. The effect of school on overweight in childhood: gain in body mass index during the school year and during summer vacation. *Am J Public Health.* 2007;97(4):696–702
  23. Roche AF, Sun SS. *Human Growth: Assessment and Interpretation.* Cambridge, United Kingdom: Cambridge University Press; 2003
  24. Ahmed ML, Yudkin PL, Macfarlane JA, McPherson K, Dunger DB. Are measurements of height made by health visitors sufficiently accurate for routine screening of growth? *Arch Dis Child.* 1990;65(12):1345–1348
  25. Cotterill AM, Majrowski WH, Hearn SJ, Jenkins S, Savage MO. Assessment of the reliability of school nurse height measurements in an inner-city population. (The Hackney Growth Initiative). *Child Care Health Dev.* 1993;19(3):159–165
  26. Lipman TH, Hench KD, Benyi T, et al. A multicentre randomised controlled trial of an intervention to improve the accuracy of linear growth measurement. *Arch Dis Child.* 2004;89(4):342–346
  27. Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr.* 1999;82(3):165–177
  28. Himes JH, Bouchard C. Validity of anthropometry in classifying youths as obese. *Int J Obes.* 1989;13(2):183–193
  29. Barlow S; Expert Committee. Expert Committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report. *Pediatrics.* 2007;120(suppl 4):S164–192
  30. Must A, Dallal GE, Dietz WH. Reference data for obesity: 85th and 95th percentiles of body mass index (wt/ht<sup>2</sup>) and triceps skinfold thickness. *Am J Clin Nutr.* 1991;53(4):839–846
  31. Bellizzi MC, Dietz WH. Workshop on childhood obesity: summary of the discussion. *Am J Clin Nutr.* 1999;70(1 pt 2):173S–175S
  32. World Health Organization. *Obesity: Preventing and Managing the Global Epidemic—Report of a WHO Consultation on Obesity, Geneva, 3–5 June 1997.* Geneva, Switzerland: World Health Organization; 1998
  33. Li M, Dibley MJ, Sibbritt D, Yan H. An assessment of adolescent overweight and obesity in Xi'an City, China. *Int J Pediatr Obes.* 2006;1(1):50–58
  34. Laxmaiah A, Nagalla B, Vijayaraghavan K, Nair M. Factors affecting prevalence of overweight among 12- to 17-year-old urban adolescents in Hyderabad, India. *Obesity (Silver Spring).* 2007; 15(6):1384–1390
  35. de Onis M, Garza C, Onyango AW, Martorell R. WHO child growth standards. *Acta Paediatr.* 2006;95(suppl 450):1–101
  36. de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martinez J. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. *Food Nutr Bull.* 2004;25(1):S15–26
  37. de Onis M, Garza C, Onyango AW, Borghi E. Comparison of the WHO child growth standards and the CDC 2000 growth charts. *J Nutr.* 2007;137(1):144–148
  38. Mei Z, Ogden CL, Flegal KM, Grummer-Strawn LM. Comparison of the prevalence of shortness, underweight, and overweight among US children aged 0 to 59 months by using the CDC 2000 and the WHO 2006 growth charts. *J Pediatr.* 2008;153(5):622–628
  39. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ.* 2007;85(9): 660–667
  40. Hamill PV, Drizd TA, Johnson CL, Reed RB, Roche AF. NCHS growth curves for children birth-18 years: United States. *Vital Health Stat 11.* 1977;(165):i–iv, 1–74
  41. Himes JH. Agreement among anthropometric indicators identifying the fattest adolescents. *Int J Obes Relat Metab Disord.* 1999;23(suppl 2):S18–S21

42. Krebs NF, Jacobson MS; American Academy of Pediatrics, Committee on Nutrition. Prevention of pediatric overweight and obesity. *Pediatrics*. 2003;112(2):424–430
43. Hedley AA, Ogden CL, Johnson CL, Carroll MD, Curtin LR, Flegal KM. Prevalence of overweight and obesity among US children, adolescents, and adults, 1999–2002. *JAMA*. 2004;291(23):2847–2850
44. Koplan J, Liverman C, Kraak V. *Preventing Childhood Obesity: Health in the Balance*. Washington, DC: National Academies Press; 2005
45. Freedman DS, Wang J, Maynard LM, et al. Relation of BMI to fat and fat-free mass among children and adolescents. *Int J Obes (Lond)*. 2005;29(1):1–8
46. Mei Z, Grummer-Strawn LM, Wang J, et al. Do skinfold measurements provide additional information to body mass index in the assessment of body fatness among children and adolescents? *Pediatrics*. 2007;119(6). Available at: [www.pediatrics.org/cgi/content/full/119/6/e1306](http://www.pediatrics.org/cgi/content/full/119/6/e1306)
47. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*. 2000;320(7244):1240–1243
48. National Institutes of Health. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report [published correction appears in *Obes Res*. 1998;6(6):464]. *Obes Res*. 1998;6(suppl 2):51S–209S
49. Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc*. 1927;22(158):202–212
50. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med*. 1998;17(8):857–872
51. Freedman DS, Wang J, Ogden CL, et al. The prediction of body fatness by BMI and skinfold thicknesses among children and adolescents. *Ann Hum Biol*. 2007;34(2):183–194
52. Guo SS, Roche AF, Chumlea WC, Johnson C, Kuczmarski RJ, Curtin R. Statistical effects of varying sample sizes on the precision of percentile estimates. *Am J Hum Biol*. 2000;12(1):64–74
53. World Health Organization. *Physical Status: The Use and Interpretation of Anthropometry—Report of a WHO Expert Committee*. Geneva, Switzerland: World Health Organization; 1995. WHO technical report series 854
54. Styne D. Growth. In: Greenspan FS, Gardner DG, eds. *Basic and Clinical Endocrinology*. 6th ed. New York, NY: Lange Medical Books/McGraw Hill; 2001:163–200
55. Gordon CC, Chumlea WC, Roche AF. Stature, recumbent length, and weight. In: Lohman TG, Roche AF, Martorell R, eds. *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics Books; 1988:3–8
56. Zeman J, Cassano M, Perry-Parrish C, Stegall S. Emotion regulation in children and adolescents. *J Dev Behav Pediatr*. 2006;27(2):155–168
57. Lohman T, Roche A, Martorell R, eds. *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics Books; 1988
58. Cameron N. *The Measurement of Human Growth*. London, United Kingdom: Croom Helm; 1984
59. Habicht JP. Standardization of quantitative epidemiological methods in the field [in Spanish]. *Bol Oficina Sanit Panam*. 1974;76(5):375–384
60. Tate RF, Klett GW. Optimal confidence intervals for the variance of a normal distribution. *J Am Stat Assoc*. 1959;54(287):674–682
61. Himes JH. Reliability of anthropometric methods and replicate measurements. *Am J Phys Anthropol*. 1989;79(1):77–80
62. Sherry B, Jefferds ME, Grummer-Strawn LM. Accuracy of adolescent self-report of height and weight in assessing overweight status: a literature review. *Arch Pediatr Adolesc Med*. 2007;161(12):1154–1161
63. Davis H, Gergen PJ. Mexican-American mothers' reports of the weights and heights of children 6 months through 11 years old. *J Am Diet Assoc*. 1994;94(5):512–516
64. Himes JH, Faricy A. Validity and reliability of self-reported stature and weight of US adolescents. *Am J Hum Biol*. 2001;13(2):255–260
65. Himes JH, Story M. Validity of self-reported weight and stature of American Indian youth. *J Adolesc Health*. 1992;13(2):118–120
66. Hauck FR, White L, Cao G, Woolf N, Strauss K. Inaccuracy of self-reported weights and heights among American Indian adolescents. *Ann Epidemiol*. 1995;5(5):386–392
67. Brenner ND, McManus T, Galuska DA, Lowry R, Wechsler H. Reliability and validity of self-reported height and weight among high school students. *J Adolesc Health*. 2003;32(4):281–287
68. Himes JH, Hannan P, Wall M, Neumark-Sztainer D. Factors associated with errors in self-reports of stature, weight, and body mass index in Minnesota adolescents. *Ann Epidemiol*. 2005;15(4):272–278

69. Tsigilis N. Can secondary school students' self-reported measures of height and weight be trusted? An effect size approach. *Eur J Public Health*. 2006;16(5):532–535
70. Brooks-Gunn J, Warren MP, Rosso J, Gargiulo J. Validity of self-report measures of girls' pubertal status. *Child Dev*. 1987;58(3):829–841
71. Shannon B, Smiciklas-Wright H, Wang MQ. Inaccuracies in self-reported weights and heights of a sample of sixth-grade children. *J Am Diet Assoc*. 1991;91(6):675–678
72. Jansen W, van de Looij-Jansen PM, Ferreira I, de Wilde EJ, Brug J. Differences in measured and self-reported height and weight in Dutch adolescents. *Ann Nutr Metab*. 2006;50(4):339–346
73. Kolbe LJ, Kann L, Collins JL. Overview of the Youth Risk Behavior Surveillance System. *Public Health Rep*. 1993;108(suppl 1):2–10
74. Neumark-Sztainer D, Wall M, Haines J, Story M, Eisenberg ME. Why does dieting predict weight gain in adolescents? Findings from project EAT-II: a 5-year longitudinal study. *J Am Diet Assoc*. 2007;107(3):448–455
75. Neumark-Sztainer D, Story M, Hannan PJ, Perry CL, Irving LM. Weight-related concerns and behaviors among overweight and nonoverweight adolescents: implications for preventing weight-related disorders. *Arch Pediatr Adolesc Med*. 2002;156(2):171–178
76. Wing RR, Epstein LH, Neff D. Accuracy of parents' reports of height and weight. *J Behav Assess*. 1980;2(2):105–110
77. Sekine M, Yamagami T, Hamanishi S, Kagamimori S. Accuracy of the estimated prevalence of childhood obesity from height and weight values reported by parents: results of the Toyama Birth Cohort study. *J Epidemiol*. 2002;12(1):9–13
78. Huybrechts I, De Bacquer D, Van Trimpont I, De Backer G, De Henauw S. Validity of parentally reported weight and height for preschool-aged children in Belgium and its impact on classification into body mass index categories. *Pediatrics*. 2006;118(5):2109–2118
79. Dubois L, Girard M. Accuracy of maternal reports of pre-schoolers' weights and heights as estimates of BMI values. *Int J Epidemiol*. 2007;36(1):132–138
80. Scholtens S, Brunekreef B, Visscher TL, et al. Reported versus measured body weight and height of 4-year-old children and the prevalence of overweight. *Eur J Public Health*. 2007;17(4):369–374
81. Garcia-Marcos L, Valverde-Molina J, Sanchez-Solis M, et al. Validity of parent-reported height and weight for defining obesity among asthmatic and nonasthmatic schoolchildren. *Int Arch Allergy Immunol*. 2006;139(2):139–145
82. Akerman A, Williams ME, Meunier J. Perception versus reality: an exploration of children's measured body mass in relation to caregivers' estimates. *J Health Psychol*. 2007;12(6):871–882
83. Harrison GG, Buskirk ER, Carter JEL, et al. Skinfold thicknesses and measurement technique. In: Lohman TG, Roche AF, Martorell R, eds. *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics Books; 1988:55–70
84. Himes J. Subcutaneous fat thickness as an indicator of nutritional status. In: Green L, Johnston P, eds. *Social and Biological Predictors of Nutritional Status, Physical Growth, and Neurological Development*. New York, NY: Academic Press; 1980:9–32
85. Norgan NG. Anthropometric assessment of body fat and fatness. In: Himes JH, ed. *Anthropometric Assessment of Nutritional Status*. New York, NY: Wiley-Liss; 1991:197–212
86. Roche AF, Sievogel RM, Chumlea WC, Webb P. Grading body fatness from limited anthropometric data. *Am J Clin Nutr*. 1981;34(12):2831–2838
87. Sarria A, Moreno LA, Garcia-Llop LA, Fleta J, Morellon MP, Bueno M. Body mass index, triceps skinfold and waist circumference in screening for adiposity in male children and adolescents. *Acta Paediatr*. 2001;90(4):387–392
88. Bedogni G, Lughetti L, Ferrari M, et al. Sensitivity and specificity of body mass index and skinfold thicknesses in detecting excess adiposity in children aged 8–12 years. *Ann Hum Biol*. 2003;30(2):132–139
89. Katzmarzyk PT, Srinivasan SR, Chen W, Malina RM, Bouchard C, Berenson GS. Body mass index, waist circumference, and clustering of cardiovascular disease risk factors in a biracial sample of children and adolescents. *Pediatrics*. 2004;114(2). Available at: [www.pediatrics.org/cgi/content/full/114/2/e198](http://www.pediatrics.org/cgi/content/full/114/2/e198)
90. Gower BA, Nagy TR, Goran MI. Visceral fat, insulin sensitivity, and lipids in prepubertal children. *Diabetes*. 1999;48(8):1515–1521
91. Lee S, Bacha F, Gungor N, Arslanian SA. Waist circumference is an independent predictor of insulin resistance in black and white youths. *J Pediatr*. 2006;148(2):188–194
92. Maffei C, Pietrobelli A, Grezzani A, Provera S, Tato L. Waist circumference and cardiovascular risk factors in prepubertal children. *Obes Res*. 2001;9(3):179–187
93. Savva SC, Tornaritis M, Savva ME, et al. Waist circumference and waist-to-height ratio are better

- predictors of cardiovascular disease risk factors in children than body mass index. *Int J Obes Relat Metab Disord.* 2000;24(11):1453–1458
94. Kahn HS, Imperatore G, Cheng YJ. A population-based comparison of BMI percentiles and waist-to-height ratio for identifying cardiovascular risk in youth. *J Pediatr.* 2005;146(4):482–488
  95. Fernández JR, Redden DT, Pietrobelli A, Allison DB. Waist circumference percentiles in nationally representative samples of African-American, European-American, and Mexican-American children and adolescents. *J Pediatr.* 2004;145(4):439–444
  96. Kyle UG, Bosaeus I, De Lorenzo AD, et al. Bioelectrical impedance analysis: part I—review of principles and methods. *Clin Nutr.* 2004;23(5):1226–1243
  97. Heyward VH, Wagner DR. *Applied Body Composition Assessment.* 2nd ed. Champaign, IL: Human Kinetics; 2004
  98. Kettaneh A, Heude B, Lommez A, Borys JM, Ducimetiere P, Charles MA. Reliability of bioimpedance analysis compared with other adiposity measurements in children: the FLVS II Study. *Diabetes Metab.* 2005;31(6):534–541
  99. Williams J, Wake M, Campbell M. Comparing estimates of body fat in children using published bioelectrical impedance analysis equations. *Int J Pediatr Obes.* 2007;2(3):174–179
  100. Lazzar S, Boirie Y, Meyer M, Vermorel M. Evaluation of two foot-to-foot bioelectrical impedance analysers to assess body composition in overweight and obese adolescents. *Br J Nutr.* 2003;90(5):987–992
  101. Chumlea WC, Guo SS, Kuczmarski RJ, et al. Body composition estimates from NHANES III bioelectrical impedance data. *Int J Obes Relat Metab Disord.* 2002;26(12):1596–1609
  102. Heymsfield SB, Lohman TG, Wang ZM, Going SB, eds. *Human Body Composition.* 2nd ed. Champaign, IL: Human Kinetics; 2005



## Challenges of Accurately Measuring and Using BMI and Other Indicators of Obesity in Children

John H. Himes

*Pediatrics* 2009;124;s3

DOI: 10.1542/peds.2008-3586D

<b>Updated Information &amp; Services</b>	including high resolution figures, can be found at: <a href="http://pediatrics.aappublications.org/content/124/Supplement_1/s3">http://pediatrics.aappublications.org/content/124/Supplement_1/s3</a>
<b>References</b>	This article cites 80 articles, 17 of which you can access for free at: <a href="http://pediatrics.aappublications.org/content/124/Supplement_1/s3#BIBL">http://pediatrics.aappublications.org/content/124/Supplement_1/s3#BIBL</a>
<b>Subspecialty Collections</b>	This article, along with others on similar topics, appears in the following collection(s): <b>Endocrinology</b> <a href="http://www.aappublications.org/cgi/collection/endocrinology_sub">http://www.aappublications.org/cgi/collection/endocrinology_sub</a> <b>Chapters Views &amp; News</b> <a href="http://www.aappublications.org/cgi/collection/chapters_views_news">http://www.aappublications.org/cgi/collection/chapters_views_news</a> <b>Obesity</b> <a href="http://www.aappublications.org/cgi/collection/obesity_new_sub">http://www.aappublications.org/cgi/collection/obesity_new_sub</a>
<b>Permissions &amp; Licensing</b>	Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: <a href="http://www.aappublications.org/site/misc/Permissions.xhtml">http://www.aappublications.org/site/misc/Permissions.xhtml</a>
<b>Reprints</b>	Information about ordering reprints can be found online: <a href="http://www.aappublications.org/site/misc/reprints.xhtml">http://www.aappublications.org/site/misc/reprints.xhtml</a>

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™



# PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

## **Challenges of Accurately Measuring and Using BMI and Other Indicators of Obesity in Children**

John H. Himes

*Pediatrics* 2009;124;s3

DOI: 10.1542/peds.2008-3586D

The online version of this article, along with updated information and services, is located on the World Wide Web at:

[http://pediatrics.aappublications.org/content/124/Supplement\\_1/s3](http://pediatrics.aappublications.org/content/124/Supplement_1/s3)

Pediatrics is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. Pediatrics is owned, published, and trademarked by the American Academy of Pediatrics, 141 Northwest Point Boulevard, Elk Grove Village, Illinois, 60007. Copyright © 2009 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 1073-0397.

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™

