American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2015. This work is written by (a) US Government employee(s) and is in the public domain in the US.

Vol. 182, No. 4
DOI: 10.1093/aje/kwv057
Advance Access publication:
July 15, 2015

Practice of Epidemiology

Comparing Methods for Identifying Biologically Implausible Values in Height, Weight, and Body Mass Index Among Youth

Hannah G. Lawman*, Cynthia L. Ogden, Sandra Hassink, Giridhar Mallya, Stephanie Vander Veur, and Gary D. Foster

* Correspondence to Dr. Hannah G. Lawman, Division of Health and Nutrition Examination Surveys, National Center for Health Statistics, Centers for Disease Control and Prevention, 3311 Toledo Road, Hyattsville, MD 20782 (e-mail: hlawman@cdc.gov).

Initially submitted January 7, 2015; accepted for publication February 25, 2015.

As more epidemiologic data on childhood obesity become available, researchers are faced with decisions regarding how to determine biologically implausible values (BIVs) in height, weight, and body mass index. The purpose of the current study was 1) to track how often large, epidemiologic studies address BIVs, 2) to review BIV identification methods, and 3) to apply those methods to a large data set of youth to determine the effects on obesity and BIV prevalence estimates. Studies with large samples of anthropometric data ($n > 1{,}000$) were reviewed to track whether and how BIVs were defined. Identified methods were then applied to a longitudinal sample of 13,662 students (65% African American, 52% male) in 55 urban, low-income schools that enroll students from kindergarten through eighth grade (ages 5–13 years) in Philadelphia, Pennsylvania, during 2011–2012. Using measured weight and height at baseline and 1-year follow-up, we compared descriptive statistics, weight status prevalence, and BIV prevalence estimates. Eleven different BIV methods were identified. When these methods were applied to a large data set, severe obesity and BIV prevalence ranged from 7.2% to 8.6% and from 0.04% to 1.68%, respectively. Approximately 41% of large epidemiologic studies did not address BIV identification, and existing identification methods varied considerably. Increased standardization of the identification and treatment of BIVs may aid in the comparability of study results and accurate monitoring of obesity trends.

biologically implausible values; body mass index; obesity; youth

Abbreviations: BIV, biologically implausible value; BMI, body mass index; CDC, Centers for Disease Control and Prevention; GHP, Get Healthy Philly.

The availability of epidemiologic data on childhood obesity has increased in response to public health concerns regarding childhood obesity trends. For example, researchers in several nations have begun to collect school surveillance data to monitor the prevalence of childhood obesity (1–5). Data quality can be variable (6), and researchers investigating obesity must address how to identify biologically implausible values (BIVs) in anthropometric data (height, weight, body mass index (BMI)). This may be of particular concern because height and weight measurements often are collected in non–research settings such as schools and when summarizing very high BMI values. In addition, a unique set of criteria may be needed for longitudinal studies to address biologically implausible change over time, such as weight loss/gain and shrinking/large height changes. The methodological process of identifying BIVs, 1 possible source of measurement error, has not been explored, and researchers have utilized a variety of methods that have not been compared or evaluated. It is not yet known how different methods for the identification of BIVs may affect study results or how often the issue of BIVs is addressed in large epidemiologic studies. Appropriate identification of BIVs will help to minimize measurement error, particularly data-recording and -entry errors, which may otherwise be more difficult to identify (6). Therefore, the current study aimed 1) to track how often large epidemiologic studies in youth address the issue of BIVs, 2) to review available methods for addressing cross-sectional and longitudinal BIVs in youth anthropometric data, and 3) to apply each unique set of BIV criteria to a large longitudinal data set from Get Healthy Philly of low-income minority youth and compare the effects on weight status and BIV prevalence estimates.

METHODS

Defining BIVs

BIV identification methods can be grouped into 3 types: externally defined standards that are based on normative reference samples, internally defined standards that are based on sample-specific criteria, and combined methods using both internally and externally defined standards.

Externally defined BIV limits. These methods rely on making comparisons of study sample anthropometric data with sources of data obtained outside the study sample (i.e., national growth charts (7–9) and age- and sex-specific ranges (10)) in order to determine if a value is biologically implausible. Other externally defined methods could include comparisons with values obtained in previously published studies (11) or with raw data found in national surveys. Some researchers have acknowledged that BMI values considered implausible under some externally defined criteria actually have been documented in adolescent bariatric populations (11). Externally defined BIV criteria are widespread, are easy to use, and create an opportunity for standardization. However, externally defined criteria are often not uniformly applied (12–14) and may not account for the heterogeneity of study samples, which could lead to the exclusion of data that are valid.

Internally defined BIV limits. Another way to define BIVs is based on the characteristics of the sample itself. This could include limits based on sample means and variances (15, 16), individually reviewed cases (17), or the consistency of BIVs with other data available for that individual (18). For example, in some national surveys, weight and height BIVs are evaluated on the basis of consistency with other available sample anthropometric measures (e.g., dual energy x-ray absorption, waist circumference, demographics). Internally defined limits allow researchers to take into account unique sample differences and all available data. These methods, however, are limited in the generalizability to other samples and hinder cross-study comparisons.

Combined methods. BIVs may also be identified by using some combination of internally and externally defined criteria. This could include the use of externally defined limits as flags (vs. fixed exclusion criteria) followed by individual review, or it could include the modification of externally referenced criteria (e.g., cutpoints from externally defined limits) (3) to fit a specific study sample. For example, researchers may select only height, weight, or BMI external BIV limits (or any combination) and provide supplemental BIV limits based on other sample characteristics or data (17). Combining and modifying methods allows researchers to take into account unique sample differences when identifying BIVs. However, this practice has led to the formation of largely unique BIV identification practices across individual studies, even among studies that reference the same BIV method (2, 12, 19–26).

Methods for identification of longitudinal change BIVs. In addition to the possibility of individual values being described as BIVs, the use of longitudinal data introduces the possibility of implausible change over time in height, weight, and BMI (e.g., extreme growth spurts, shrinkage, implausible weight increases/decreases). Longitudinal studies have the added benefit of multiple data points for an individual, which limits the usefulness of strict external BIV limits. Rather, some longitudinal studies have used methods whereby pediatric endocrinologists examine cases flagged by external limits and make case-by-case exclusions (17). This may be feasible even in very large samples because of the relatively small number of BIVs. Longitudinal data on how youths' individual body weights may plausibly change over time are lacking (27), although some research has examined normal height velocity in youth (28). Mixed and internally defined limits for identifying longitudinal change BIVs capitalize on the availability of all of an individual's longitudinal data; however, these approaches also limit generalizability and reproducibility.

Identifying BIV methods and tracking BIV awareness

In order to track how often BIV identification was discussed and to review available BIV methods, studies were identified by using PubMed, PsycInfo, and Google Scholar with search terms including "BIV," "biologically implausible," "BMI," "height," "weight," and "childhood obesity." References were cross-checked to identify additional sources such as published studies, data-processing reports, and technical manuals. Studies were included in the review if they 1) had large samples ($n > 1,000$) of youth (aged 2–18 years) with measured or self-reported height, weight, or BMI as a primary outcome or independent variable and 2) were published during or after the year 2000. These sources were assessed on how height, weight, and BMI BIVs were defined. Large samples were chosen as they were considered to have sufficient representation of BIVs, to have needed rules guiding the handling of BIVs, and to have provided an extant childhood obesity literature manageable in size to review. It is important to note that BIV decisions will likely impact smaller samples to a greater extent as the exclusion of a few data points can significantly impact results. Studies after 2000 were chosen to be consistent with when new Centers for Disease Control and Prevention (CDC) growth charts were released (7). For the purposes of tracking how often these studies addressed BIV issues, any source related to a study (e.g., technical manual, website, peer-reviewed papers) was credited for including BIV information. Multiple peer-reviewed papers reporting on the same data set (e.g., national surveys) were counted as 1 study unless the papers explicitly described different methods of identifying BIVs compared with the "parent study" documentation. Large data sets collected in regular and/or ongoing intervals, such as school surveillance data and national surveys, were considered as 1 parent study. This process revealed 42 large-scale parent studies and 11 different BIV methods. These studies and supporting materials were reviewed and grouped into 1 of 3 categories: 1) no BIV information given (in any source), 2) insufficient BIV information (including studies that provided either a BIV method, which in some cases was only partially described, or the BIV prevalence but not both), and 3) complete BIV information (including studies that provided both the complete BIV identification method and an indication of BIV prevalence).

BIV comparisons in Get Healthy Philly

Data. Existing BIV methods were applied to a large, longitudinal data set to compare the effects on relative weight outcomes and BIV prevalence. Data were from 13,662 low-income and minority youth from an urban school district with a baseline assessment in the first to sixth grade and a 1-year follow-up in the second to seventh grade. Students were from 55 schools that enroll students from kindergarten through eighth grade (ages 5–13 years) in Philadelphia, Pennsylvania (age range, 5.2–15.4 years), and data were collected as part of the Get Healthy Philly (GHP) initiative (29). The 55 schools had 94.6% (range, 73.9%–98.6%) of students eligible for free or reduced lunch. Youth were enrolled by using passive consent (consent forms sent home with students and returned only if parents did not want their child to participate). School-wide enrollment rates were 94% at baseline and 87% at follow-up. Study details, including Consolidated Standards of Reporting Trials (CONSORT) information and trends in relative weight, have been published previously (29). GHP data were selected for the current study on the basis of the large sample size and the collection of data by researchers using strict training protocols within a school-based BMI surveillance system. GHP also had a high proportion of data from subgroups with traditionally high levels of obesity.

Height and weight. Trained research assistants used standard protocols to measure height and weight (29). Youth were instructed to remove shoes, extra layers of clothing, and items from pockets. Height and weight were measured twice and averaged if the measurements were within 1 cm and 0.2 kg. If the differences were larger, a third measure was taken, and the 2 within the specified range were averaged. Weight status category was defined on the basis of sex-specific BMI-for-age percentile in the 2000 CDC growth charts: underweight (<5th percentile); normal (≥5th and <85th percentile); overweight (≥85th and <95th percentile); obese (100% to <120% of the 95th percentile); and severely obese (≥120% of the 95th percentile) (7, 30).

Demographics. Information on race, sex, month and year of birth, and grade level was obtained from schools. Race, based on parent self-report, was categorized as African American, Hispanic, Caucasian, Asian, and other.

The study was approved by the School District of Philadelphia (Pennsylvania) and institutional review boards at Temple University and the Philadelphia Department of Health. Data were collected between February 2011 and May 2012.

Procedures for applying BIV methods. After data were collected in the field, double data entry was used to enter data and to resolve discrepancies on the basis of hard copies of recorded values. After the data were thoroughly cleaned for data entry errors, data were flagged according to the BIV identification methods summarized in Table 1 and fully defined in Web Table 1 available at http://aje.oxfordjournals.org/ (categorized by letters A through K). Methods F and I were not applicable to GHP and were not used. Method F was a national survey that made BIV exclusions based on additional sample data that were not available in GHP (e.g., waist circumference). Method I was developed for children under 5 years of age (17). Methods A, D, E, and H needed modifications in order to be applied to GHP. Given the widespread use of method A based on World Health Organization recommendations (9), this method was also used longitudinally such that any individual baseline or follow-up flagged value was excluded. For method D, it was assumed that the largest 1% of changes was selected from a distribution of the absolute value of changes in BMI from baseline to follow-up (i.e., compared with the largest 0.5% of increases and decreases). Method E did not describe the age ranges used for determining age-specific mean reference values, so they were specified for each year (i.e., matched by year of age and sex). The 3rd and 97th percentile values were used in method H rather than the 1st and 99th percentiles in accordance with standard available CDC reference values (7). In addition, GHP participants did not have medical record data, which were used in method H to calculate median body weight for individuals (Web Table 1). Instead, age- and sex-specific 50th percentile national norm data were used (note: this method was retained as an internal BIV method for consistency with the original publication) (31). Working from the original GHP data set (no excluded BIVs), we applied each BIV method. Means, number of BIV exclusions, and weight status prevalence estimates were examined.

Table 1. Summary of Identification Methods for Biologically Implausible Values[a]

BIV Method (Reference No.)	Weight	Height	BMI	Change[b]
External				
A (8, 9)	X	X	X	Not applicable
B (10)	X	X	X	Not applicable
C (37)	None	None	X	Not applicable
Internal				
D (34)	None	None	None	X
E (16)	None	None	X	Not applicable
F (18)	X	X	X	Not applicable
Mixed				
G (3)	X	X	X	X
H (31)	X	X	None	Not assessed[c]
I (17)	X	X	X	Not assessed
J (29)	None	None	None	X
K (15)	None	None	X	Not assessed

Abbreviations: BIV, biologically implausible value; BMI, body mass index.

[a] Full criteria and specific definitions are listed in Web Table 1.

[b] "Not applicable" was noted for cross-sectional data; "not assessed" was noted where longitudinal data were obtained, but implausible change over time was not assessed.

[c] This criterion utilized longitudinal data in the determination of assessment for individual implausible values but did not assess criteria for implausible change over time.

RESULTS

Frequency of defining BIVs

Among the 42 large-scale studies identified (Web Table 2), 40.5% reported no BIV information (i.e., no mention of BIVs in any study-related documentation). An additional 26.2% of studies reported insufficient BIV information (i.e., lacked criteria, prevalence, or sufficient detail to be reproduced). Studies in this category most often provided documentation on unspecified height or weight ranges that were considered errors during field data collection and prompted data rechecking. For example, documentation may have reported that out-of-range values were coded as missing data but did not provide specific information on the ranges used or how many data were treated as missing under the criteria. The remaining 33.3% of studies reported both the specific BIV method used and the BIV prevalence.

Review of available BIV methods

Table 1 summarizes the 11 identified BIV methods, including 3 externally defined, 3 internally defined, and 5 combination methods. Nine of these methods addressed individual BIVs, and 3 of these methods described longitudinal BIVs (including 1 method that addressed both). Web Table 3 shows information on the samples used to develop each BIV method. One (31) of 11 methods included some assessment of validity when applying or developing the method (Web Table 3).

Because of the many unique BIV methods and combinations of methods, the extant literature varied greatly on the estimates of BIV prevalence that were reported. Overall, BIV prevalence rates reported in the literature ranged from 0.03% to 4.5% across studies (Web Table 4). Reported BIV prevalence using the internal methods was approximately 1% and ranged from 0.03% to 3.0% for external methods and from 0.06% to 4.5% for mixed methods (Web Table 4).

Applying BIV methods to Get Healthy Philly

Demographics and weight status. Participants from the full GHP sample had a mean age of 9.7 (standard deviation, 1.8) years at baseline and 10.5 (standard deviation, 1.8) years at follow-up. Overall, 1.6% were classified as underweight, 59.1% as normal weight, 16.6% as overweight, 14.0% as obese, and 8.6% as severely obese (29).

Cross-sectional BIV identification methods. Results from comparisons of available BIV identification methods are shown in Table 2. The number of BIV exclusions in the GHP data set ranged from 6 to 229 (0.04%–1.68% of the sample) depending on the criteria used. The majority of BIV exclusions were made for weight and BMI values rather than for height values. Means of unadjusted BMI, BMI percentile, and BMI z score remained relatively unchanged by the BIV exclusion criteria. Weight status category prevalence estimates differed slightly according to the BIV criteria used: 38.3%–39.2% for overweight, 21.4%–22.6% for obese, and 7.3%–8.6% for severely obese.

Longitudinal BIV identification methods. Results across the 4 longitudinal BIV identification methods are shown in Table 3. The number of BIV exclusions in the GHP data set ranged from 41 to 280 (0.3%–2.1% of the sample) depending on the criteria used. The means of unadjusted BMI, BMI percentile, and BMI z score remained relatively unchanged by the BIV exclusion criteria. Weight status category prevalence estimates differed slightly according to the

Table 2. BIVs and Weight Status by Cross-Sectional BIV Identification Method ($n = 13,662$), Get Healthy Philly Sample, 2011–2012[a]

BIV Method (Reference No.)	BIV		BIV Exclusion Criteria, no.						BMI, mean			Weight Status Category Prevalence, %		
	Cases, no.[b]	Prevalence, %	Height	Weight	BMI	Height Only[c]	Weight Only[c]	BMI Only[c]	Units[d]	Percentile	z Score	≥85th Percentile[e]	≥95th Percentile[e]	≥120% of the 95th Percentile[e]
No BIV exclusions	0	0	0	0	0	0	0	0	19.8	68.2	0.7	39.2	22.6	8.6
A (9)	229	1.68	89	121	110	68	36	40	19.6	67.7	0.7	38.4	21.6	7.4
B (10)	46	0.34	4	10	40	2	2	34	19.7	68.1	0.7	39.1	22.4	8.3
C (37)	53	0.38	0	0	53	0	0	53	19.7	68.1	0.7	39	22.3	8.2
E (16)	74	0.54	0	0	74	0	0	74	19.7	68	0.7	38.9	22.2	8.1
G (3)	118	0.86	2	48	110	1	6	69	19.6	67.9	0.7	38.7	21.9	7.8
H (31)	6	0.04	2	4	0	2	4	0	19.8	68.2	0.7	39.2	22.6	8.5
K (15)	199	1.46	0	0	199	0	0	199	19.5	67.7	0.7	38.3	21.5	7.2

Abbreviations: BIV, biologically implausible value; BMI, body mass index; GHP, Get Healthy Philly.

[a] Not all BIV methods were applicable to the GHP data or cross-sectional data. Method F uses individual review and comparison with waist circumference and dual-energy x-ray absorptiometry for any values exceeding the 99th percentile or below the 1st percentile in height, weight, or BMI (an internally defined comparison using data collected for national surveys). In the GHP sample, height ($n = 425$), weight ($n = 916$), and BMI ($n = 908$) values would have been marked for individual review using national data as an external defined limit.

[b] Height, weight, and BMI BIV counts may not sum to total BIV cases because some cases were identified as BIV via multiple markers.

[c] Counts only those values without any other co-occurring BIV indicator.

[d] Expressed as weight (kg)/height (m)^2.

[e] Overweight: ≥85th percentile; obese: 100% to <120% of the 95th percentile; severely obese: ≥120% of the 95th percentile.

Table 3. BIV and Weight Status by Longitudinal BIV Identification Method ($n = 13{,}662$), Get Healthy Philly Sample, 2011–2012[a]

BIV Method (Reference No.)	BIV		BIV Exclusion Criteria, no.						BMI, mean			Weight Status Category Prevalence, %		
	Cases, no.[b]	Prevalence, %	Height	Weight	BMI	Height Only[c]	Weight Only[c]	BMI Only[c]	Units[d]	Percentile	z Score	≥85th Percentile[e]	≥95th Percentile[e]	≥120% of the 95th Percentile[e]
No BIV exclusion	0	0	0	0	0	0	0	0	19.8	68.2	0.7	39.2	22.6	8.6
A (9)	280	2.05	117	144	130	92	41	44	19.5	67.7	0.7	38.3	21.4	7.3
D (34)	137	1	0	0	137	0	0	137	19.7	68.1	0.7	39	22.3	8.2
G (3)	66	0.48	6	0	62	3	0	59	19.7	68.1	0.7	39.1	22.4	8.3
J (29)	41	0.3	19	0	22	19	0	22	19.8	68.2	0.7	39.2	22.6	8.5

Abbreviations: BIV, biologically implausible values; BMI, body mass index.

[a] Not all BIV methods were applicable to the Get Healthy Philly data or to longitudinal data.

[b] Height, weight, and BMI BIV counts may not sum to total BIV cases because some cases were identified as BIV via multiple markers.

[c] Counts only those values without any other co-occurring BIV indicator.

[d] Expressed as weight (kg)/height (m)^2.

[e] Overweight: ≥85th percentile to <95th percentile; obese: 100% to <120% of the 95th percentile; severely obese: ≥120% of the 95th percentile.

BIV criteria used: 38.3%–39.2% for overweight, 21.5%–22.6% for obese, and 7.2%–8.6% for severe obesity.

DISCUSSION

The current study's aims were to assess how often large epidemiologic studies address BIV identification, to review available methods for addressing BIVs, and to examine the impact of varying BIV criteria in a large sample of youth. There were 3 key findings.

First, approximately 67% of large epidemiologic studies conducted since 2000 ($n > 1{,}000$) reported no information or insufficient information regarding the identification of BIVs. This included almost 41% of studies that did not mention BIVs in any study documentation. This percentage may be higher among studies with smaller samples. BIV identification has statistical implications that mirror those of the accurate identification of outliers and the occurrence of measurement error, including reduced power and/or biased parameter estimates (32). Furthermore, this study and others (11, 17) suggest that any bias is compounded by the fact that most BIV cases in the United States are concentrated at the upper end of the distribution, introducing the possibility of systematic error. Even so, BIV identification is relevant both for data in the tails of the distribution (e.g., outliers) and for implausible changes in the middle of the distribution (e.g., change from the 20th percentile to the 70th percentile). BIVs may also be relevant for plausible values in the middle of the distribution with implausible corresponding data (e.g., waist circumference, weight-for-age). The current study demonstrated that researchers are not consistently considering BIV identification and/or are not consistently reporting BIV-related practices.

Second, the current study found that a wide variety of methods, which have not been evaluated, were being used to identify BIVs. Only 1 (31) of the 11 identified BIV methods included some assessment of validity (i.e., proportion of valid data marked implausible as assessed with medical record data). Two methods included individual review by expert pediatric physicians. This has implications for the accurate assessment of childhood obesity in large epidemiologic studies, as well as the comparability of study findings. The wide variation in BIV identification was exemplified when considering the most widely used BIV method. These limits, originally proposed by the World Health Organization (9), have become widely available through the CDC SAS program (SAS Institute, Inc., Cary, North Carolina) for the calculation of youth anthropometric percentile and z-score values based on the 2000 CDC growth charts (7, 8). A number of modified versions of the original criteria have been used in the literature, such as the use of only height and weight limits, only BMI limits, and modification of the cutoffs (3, 17, 33). A relatively high rate of flagged BIVs using the original cutoffs may, in part, have led to inconsistent applications or modifications of the method. In the current study, the original method displayed the highest BIV prevalence compared with all the other methods applied to the data. This high rate may be a combination of incorrectly flagging valid high values, as well as flagging data errors. Previous research has concluded that some BIV methods may lead to the exclusion of valid data (11, 17). Lo et al. (17) investigated severe obesity in preschoolers using several internally and externally defined height, weight, and BMI BIV limits. The last BIV identification step included individual review of implausible values by a pediatric endocrinologist and ultimately excluded 31 of 430 values (17).

Other methods applied to GHP showed relatively few BIVs, and they may fail to adequately identify erroneous data. It is unclear which of the existing methods, if any, are valid, and no methods have been evaluated against a "gold standard" (e.g., verified data from an independent measurement). In addition, the quality of data collection techniques varies (e.g., trained researchers, physical education teachers, single and double data entry), and BIV identification strategies may need to vary accordingly (e.g., additional checks if high BIV rate). Methods that are not reproducible or imprecisely defined methods may limit cross-study comparisons (16, 31, 34). Taken together, previous studies (11, 17) and results from the current study suggest that future research is needed on how BIVs can be accurately and systematically identified.

Third, the variation in methods for identifying BIVs impacted estimates of weight status and BIV prevalence and sample size. In the current study, severe obesity estimates ranged from 7.2% to 8.6% (with no BIV exclusions), depending on the BIV method used. Although the absolute difference in prevalence across methods is relatively small (1.4 percentage points), this variation constituted a substantial portion of the severe obesity prevalence estimate (16% of the estimate is due to the BIV method used). BIV identification may contribute to the range of severe obesity prevalence shown across studies and may differentially over- or underestimate prevalence depending on which BIV criteria are used or on which subpopulations they are applied (30, 35). For example, school surveillance data from New York, New York, excluded 0.08% of individuals in 1998 and 2%–3% of youth in 2007–2011 for BIVs, which illustrates how the impact of BIV decisions may change over time (23, 26). There may be systematic bias in the estimates of severe obesity prevalence in instances where accurate high values are routinely excluded as BIVs (or where inaccurate high values are retained). In addition, many data sets now collected arise from non-research settings, such as schools and medical records, where training and equipment are variable. A higher prevalence of severe obesity has been observed in ethnic minority groups, and some BIV methods may differentially impact severe obesity estimates across samples with different demographic characteristics.

The current study has implications for research relating to BIVs in anthropometric data. Increased standardization in the reporting of BIV issues, such as consistently reporting BIV procedure, criteria, and prevalence, could inform cross-study comparisons and aid in understanding BIVs and trends of very high BMI values. Of note, the majority of BIV methods (9 of 11) focused exclusively on extreme BIVs at the tails of the distribution. However, a variety of BIV indicators may need to be considered in order to capture BIVs both near the center and at the tails of the distribution where BIV values may be more visible. This would require longitudinal data or additional variables in the data set that can inform a BIV decision (e.g., waist circumference, body composition) and would increase the difficulty of data collection if these data are not already being collected. An added consideration with commonly used World Health Organization BIV limits (9) has arisen from the use of different methods for z-score calculations in national reference data. The method used with the 2000 CDC growth charts makes extreme z scores mathematically impossible, and thus BIVs would not be identified with the BMI z-score variable (8). Specifically, the CDC lambda-mu-sigma (LMS) statistical model uses a Box-Cox transformation to account for skewness, which results in all extreme values (plausible or implausible) being mapped to a high plausible z score (36). Although the CDC has provided alternative calculations to be used as flags for the purposes of BIV identification (8), it is unclear the extent to which researchers use these flags versus the calculated BMI z scores for BIV identification. Researchers may need to consider the impact of BIV decisions on analyses, including meaningful statistical and clinical implications or sensitivity analyses, particularly when addressing very high BMI values. Although the appropriate method for detecting BIVs is not yet clear, BIV decisions may be among the easier sources of measurement error to detect.

Strengths of the study included the comparison of multiple methods of BIV identification and the large, diverse longitudinal sample. Limitations of the current study should also be considered. The GHP data had a short follow-up period, and no gold standard for identifying BIVs was available. In addition, it was outside the scope of the paper to track the frequency of BIV identification in small samples.

The current study reviewed 11 methods for identifying BIVs. Furthermore, 41% of the large epidemiologic studies reviewed provided no information on BIVs, and an additional 26% provided insufficient information on BIV identification. The method used to identify BIVs impacts BIV prevalence and, as a result, sample size and to a small degree the estimates of severe obesity prevalence. Increased standardization of the identification and treatment of BIVs may aid in the comparability of results across studies and the accurate monitoring of childhood obesity trends.

ACKNOWLEDGMENTS

Author affiliations: Center for Obesity Research and Education, Temple University School of Medicine, Philadelphia, Pennsylvania (Hannah G. Lawman, Stephanie Vander Veur, Gary D. Foster); National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Maryland (Hannah G. Lawman, Cynthia L. Ogden); Department of Pediatrics, Nemours/Alfred I. duPont Hospital for Children, Wilmington, Delaware (Sandra Hassink); and Philadelphia Department of Public Health, Philadelphia, Pennsylvania (Giridhar Mallya).

This work was supported by a grant (cooperative agreement 3U58DP002626-01S1) from the Centers for Disease Control and Prevention and by grants (1-P60-DK020541 and 1-F32-DK100248-01) from the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health.

The findings and conclusions in this manuscript are those of the authors and not necessarily those of the Centers for Disease Control and Prevention.

While the study was being conducted, G.D.F. served on scientific advisory boards for Con Agra Foods, United Health Group, and Tate & Lyle. Currently, G.D.F. and S.V.V. are employed by Weight Watchers International. The other authors report no conflicts.

REFERENCES

1. Ali A, Becker E, Chaundhury M, et al. *Health Survey for England—2006: Methodology and Documentation*. Leeds, United Kingdom: Health & Social Care Information Centre; 2006.
2. Madsen KA, Weedn AE, Crawford PB. Disparities in peaks, plateaus, and declines in prevalence of high BMI among adolescents. *Pediatrics*. 2010;126(3):434–442.
3. Kim J, Must A, Fitzmaurice GM, et al. Incidence and remission rates of overweight among children aged 5 to 13 years in a district-wide school surveillance system. *Am J Public Health*. 2005;95(9):1588–1594.
4. Robbins JM, Mallya G, Polansky M, et al. Prevalence, disparities, and trends in obesity and severe obesity among students in the Philadelphia, Pennsylvania, School District, 2006–2010. *Prev Chronic Dis*. 2012;9:E145.
5. Savva SC, Kourides Y, Tornaritis M, et al. Obesity in children and adolescents in Cyprus. Prevalence and predisposing factors. *Int J Obes Relat Metab Disord*. 2002;26(8):1036–1045.
6. Himes JH. Challenges of accurately measuring and using BMI and other indicators of obesity in children. *Pediatrics*. 2009;124(suppl 1):S3–S22.
7. Kuczmarski RJ, Ogden CL, Guo SS, et al. 2000 CDC growth charts for the United States: methods and development. *Vital Health Stat 11*. 2002;(246):1–190.
8. Centers for Disease Control and Prevention. Cut-offs to define outliers in the 2000 CDC growth charts. http://www.cdc.gov/nccdphp/dnpa/growthcharts/resources/BIV-cutoffs.pdf. Updated October 24, 2014. Accessed February 24, 2015.
9. Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. *World Health Organ Tech Rep Ser*. 1995;854:1–452.
10. Youth Risk Behavior Surveillance System. 2013 YRBS data user's guide. ftp://ftp.cdc.gov/pub/data/yrbs/2011/YRBS_2011_National_User_Guide.pdf. Published June 2012. Accessed February 24, 2015.
11. Ratcliff MB, Jenkins TM, Reiter-Purtill J, et al. Risk-taking behaviors of adolescents with extreme obesity: normative or not? *Pediatrics*. 2011;127(5):827–834.
12. Novotny R, Oshiro CES, Wilkens LR. Prevalence of childhood obesity among young multiethnic children from a health maintenance organization in Hawaii. *Child Obes*. 2013;9(1):35–42.
13. Berlan ED, Corliss HL, Field AE, et al. Sexual orientation and bullying among adolescents in the Growing Up Today Study. *J Adolesc Health*. 2010;46(4):366–371.
14. Bogen DL, Hanusa BH, Whitaker RC. The effect of breast-feeding with and without formula use on the risk of obesity at 4 years of age. *Obes Res*. 2004;12(9):1527–1535.
15. Field AE, Austin SB, Taylor CB, et al. Relation between dieting and weight change among preadolescents and adolescents. *Pediatrics*. 2003;112(4):900–906.
16. Conde WL, Monteiro CA. Body mass index cutoff points for evaluation of nutritional status in Brazilian children and adolescents. *J Pediatr (Rio J)*. 2006;82(4):266–272.
17. Lo JC, Maring B, Chandra M, et al. Prevalence of obesity and extreme obesity in children aged 3-5 years. *Pediatr Obes*. 2014;9(3):167–175.
18. National Health and Nutrition Examination Survey. 2001–2002 data documentation, codebook, and frequencies: body measurements. http://www.cdc.gov/nchs/nhanes/nhanes2001-2002/BMX_B.htm. Published May 2004. Accessed February 24, 2015.
19. Gundersen C, Lohman BJ, Eisenmann JC, et al. Child-specific food insecurity and overweight are not associated in a sample of 10- to 15-year-old low-income youth. *J Nutr*. 2008;131(2):371–378.
20. Nader PR, O'Brien M, Houts R, et al. Identifying risk for obesity in early childhood. *Pediatrics*. 2006;118(3):e594–e601.
21. Mei Z, Scanlon KS, Grummer-Strawn LM, et al. Increasing prevalence of overweight among US low-income preschool children: the Centers for Disease Control and Prevention Pediatric Nutrition Surveillance, 1983 to 1995. *Pediatrics*. 1998;101(1):E12.
22. Weedn AE, Ang SC, Zeman CL, et al. Obesity prevalence in low-income preschool children in Oklahoma. *Clin Pediatr (Phila)*. 2012;51(10):917–922.
23. Melnik TA, Rhoades SJ, Wales KR, et al. Overweight school children in New York City: prevalence estimates and characteristics. *Int J Obes Relat Metab Disord*. 1998;22(1):7–13.
24. Sekhobo JP, Edmunds LS, Reynolds DK, et al. Trends in prevalence of obesity and overweight among children enrolled in the New York State WIC Program, 2002–2007. *Public Health Rep*. 2010;125(2):218–224.
25. Pan L, Blanck HM, Sherry B, et al. Trends in the prevalence of extreme obesity among US preschool-aged children living in low-income families, 1998–2010. *JAMA*. 2012;308(24):2563–2565.
26. Berger M, Konty K, Day S, et al. Obesity in K-8 students—New York City, 2006–07 to 2010–11 school years. *Morb Mortal Wkly Rep*. 2011;60(49):1673–1678.
27. Zemel BS. A commentary on the construction of weight velocity charts. *Nutr Clin Pract*. 2009;24(5):651–653.
28. Tanner JM, Davies PS. Clinical longitudinal standards for height and height velocity for North American children. *J Pediatr*. 1985;107(3):317–329.
29. Lawman HG, Mallya G, Veur SV, et al. Trends in relative weight over 1 year in low-income urban youth. *Obesity (Silver Spring)*. 2015;23(2):436–442.
30. Flegal KM, Wei R, Ogden CL, et al. Characterizing extreme values of body mass index-for-age by using the 2000 Centers for Disease Control and Prevention growth charts. *Am J Clin Nutr*. 2009;90(5):1314–1320.
31. Smith N, Coleman KJ, Lawrence JM, et al. Body weight and height data in electronic medical records of children. *Int J Pediatr Obes*. 2010;5(3):237–242.
32. Cohen J, Cohen P, West SG, et al. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2003.
33. Dennison BA, Edmunds LS, Stratton HH, et al. Rapid infant weight gain predicts childhood overweight. *Obesity (Silver Spring)*. 2006;14(3):491–499.
34. Sturm R, Datar A. Body mass index in elementary school children, metropolitan area food prices and food outlet density. *Public Health*. 2005;119(12):1059–1068.
35. Koebnick C, Smith N, Coleman KJ, et al. Prevalence of extreme obesity in a multiethnic cohort of children and adolescents. *J Pediatr*. 2010;157(1):26–31.e2.
36. Woo JG. Using body mass index *z*-score among severely obese adolescents: a cautionary note. *Int J Pediatr Obes*. 2009;4(4):405–410.
37. Lobstein TJ, James WP, Cole TJ. Increasing levels of excess weight among children in England. *Int J Obes Relat Metab Disord*. 2003;27(9):1136–1138.