**Objective:** To examine the accuracy of self-reported height and weight data to classify adolescent overweight status. Self-reported height and weight are commonly used with minimal consideration of accuracy.

**Data Sources:** Eleven studies (4 nationally representative, 7 convenience sample or locally based).

**Study Selection:** Peer-reviewed articles of studies conducted in the United States that compared self-reported and directly measured height, weight, and/or body mass index data to classify overweight among adolescents.

**Main Exposures:** Self-reported and directly measured height and weight.

**Main Outcome Measures:** Overweight prevalence; missing data, bias, and accuracy.

**Results:** Studies varied in examination of bias. Sensitivity of self-reported data for classification of overweight ranged from 55% to 76% (4 of 4 studies). Overweight prevalence was −0.4% to −17.7% lower when body mass index was based on self-reported data vs directly measured data (5 of 5 studies). Females underestimated weight more than males (ranges, −4.0 to −1.0 kg vs −2.6 to 1.5 kg, respectively) (9 of 9 studies); overweight individuals underestimated weight more than nonoverweight individuals (6 of 6 studies). Missing self-reported data ranged from 0% to 23% (9 of 9 studies). There was inadequate information on bias by age and race/ethnicity.

**Conclusions:** Self-reported data are valuable if the only source of data. However, self-reported data underestimate overweight prevalence and there is bias by sex and weight status. Lower sensitivities of self-reported data indicate that one-fourth to one-half of those overweight would be missed. Other potential biases in self-reported data, such as across subgroups, need further clarification. The feasibility of collecting directly measured height and weight data on a state/community level should be explored because directly measured data are more accurate.

Arch Pediatr Adolesc Med. 2007;161(12):1154-1161

---

One important aspect of addressing pediatric overweight as a national public health priority is to document and monitor the prevalence. Prevalence data describe the extent of the problem and show trends over time; they are also useful for targeting interventions to geographic areas or subgroups with the highest prevalence. Monitoring changes in prevalence is an important component of evaluating intervention programs to prevent and treat overweight; in particular, these data are useful for understanding and guiding program effectiveness, improvement, and maintenance. The most commonly used anthropometric screening tool for overweight is body mass index (BMI) (calculated as weight in kilograms divided by height in meters squared). Thus, height and weight data are essential for classifying overweight status. How these data are collected becomes an important issue as state health departments and researchers more actively try, with limited resources, to monitor the impact of programs to reduce the obesity epidemic.

Self-reported height and weight are sometimes used to determine and/or monitor overweight. The self-reported method is chosen in preference to direct measurement because self-reported data are easier, cheaper, and quicker to collect. Moreover, direct measurement of height and weight requires appropriate training and monitoring of personnel and the use of accurate, well-maintained equipment. Because of limited resources, self-reported data may be the only data available. Unfortunately, minimal attention is given to potential inaccuracies or bias introduced with the self-reported method. Moreover, articles in the literature may give no
or limited mention of the quality of self-reported data. For example, Valenti et al. reported overweight prevalence of Australian children with no mention of the quality of their parental-reported data; Sen3 examined the frequency of family dinner and weight status based on self-reported data from the National Longitudinal Survey of Youth. Sen cites the Strauss4 National Health and Nutrition Examination Survey III analysis, which reported that 94% of adolescents had their weight status correctly classified, as justification for the accuracy of self-reported data. As will be seen in this review, a single summary statistic can mask bias.

The goals of this review were to examine and summarize the accuracy of using self-reported data in contrast to directly measured data for identifying and monitoring overweight among US adolescents. We examined consistency across studies in the magnitude of bias by age, sex, race/ethnicity, and prevalence of overweight and the extent of prevalence bias by sensitivity and specificity.

**METHODS**

**LITERATURE SEARCH**

We used a comprehensive literature search of the electronic databases PubMed and MEDLINE, including “in-process MEDLINE.” The search terms included body weight, body height, and body mass index; adolescent, youth, or teenager; self-report, self-disclosure, and self-assessment; reproducibility of results, sensitivity and specificity, analysis of variance, and multivariate analysis; valid, reliable, and accurate; and English language. We did not limit the years of our search. We identified 60 articles, but only 12 studies met our inclusion criteria of peer-reviewed articles conducted in the United States that compared self-reported and directly measured height and weight data or BMIs calculated from these different measures for adolescents. We excluded 1 additional study that did not include correlation coefficients between self-reported and directly measured data or sensitivity and specificity analyses but only presented self-reported and directly measured mean differences. Two studies used National Health and Nutrition Examination Survey III data for their analyses but examined different aspects of bias, so we included both.

**CODING AND ANALYSIS OF STUDIES**

For each study, we identified the geographic location; study design, including sampling protocol, setting, sequence of measurements, measurement protocols, definitions of at risk for overweight and/or overweight status, including reference sources and exclusion of biologically implausible values; and population characteristics, including age, race/ethnicity, sample size, and directly measured prevalence of overweight. For each study, we extracted percentage of missing data; differences between the self-reported and directly measured height, weight, and BMI; stratification of these differences by age, sex, race/ethnicity, and overweight status, when available; regression analysis to identify predictors of self-reported bias; differences between self-reported and directly measured prevalence for both at risk for overweight and overweight; correlation coefficients between self-reported and directly measured height and weight and calculated BMI; and sensitivity, specificity, and positive predictive value data for prevalence. When data were not available, we contacted authors in an attempt to obtain complete information. For our review, 2 of us (B.S. and M.E.J.) independently summarized each study and resolved differences by discussion. We analyzed the consistency of patterns across the 11 studies for the previously listed parameters.

Meta-analysis or a moderator analysis was not appropriate to define an effect size for this review because there were only 11 studies and they varied substantially in sample size and composition, study design, and definition of overweight. Several studies included small convenience samples or were locally based, making their findings nongeneralizable to the overall US population of adolescents.

**RESULTS**

An overview of the 11 studies (4 based on nationally representative samples and 7 convenience sample or locally based studies) is summarized in Table 1. We stratified our summary of studies into these 2 categories to differentiate the generalizability of their results. All studies included children 11 years or older. Human subjects’ approval was obtained for all of the 11 studies.

**DIFFERENCES IN STUDY METHODS**

Data for the 4 nationally representative studies were collected during the early 1980s through the mid 1990s; all used standardized protocols for measurements. Intervals between collection of self-reported and directly measured data ranged from the same day to a few weeks. Only Himes and Faricy described exclusion of data with biologically implausible values: they excluded subjects with differences between self-reported and directly measured height or weight values more than 5 SDs of the overall mean differences.

Among the 7 convenience sample or locally based studies, data were collected over a wider period (1970s through 2000); 3 studies did not report year(s) of data collection. Various measurement protocols were used: used referenced protocols, used standardized but un-referenced protocols, described some procedures, and did not describe. The interval between the collection of self-reported and directly measured data ranged from the same day to a month or so; one did not report the interval. Two studies described exclusions of biologically implausible values: Shannon et al. “zero” values and height values greater than 400 in, and Stewart, height and weight values 4 or more SDs from mean difference of self-reported and directly measured data.

Moreover, these 11 studies used a variety of references and/or cutoffs to define overweight status. These variations made comparisons among studies difficult.

**MISSING DATA**

Among the national surveys, missing data (unknown/unsurable) for self-reported values ranged from 14% to 37%. Himes and Faricy showed an effect by age: 41% of 12-year-olds and 25% of 13-year-olds did not report their weight, in contrast to only 2% of the 14- to 16-year-olds. Their logistic regression model showed that younger age, being male, shorter directly measured height, and lighter directly measured weight were all independently and significantly associated with missing data.
Only 5 convenience sample or locally based studies reported missing data, which ranged from 0% to 31%.10,11,13-15 Shannon et al11 found no differences in missing self-reported data by height, weight, or sex, but they did not examine the effect by age in their study of 11- to 13-year-olds.

Table 2 shows the differences in mean (self-reported–directly measured) height, weight, and BMI by sex for the 9 studies with these data available. For height, weight, and BMI, the magnitude of mean bias by sex was greater for participants in the convenience sample or locally based studies than in the nationally representative studies.

**DIFFERENCES IN MEANS**

**Height**

Height was both overestimated and underestimated. For the 7 studies for both males and females, the range of mean height differences was relatively small at −1.1 cm to 2.4 cm, with the exception of Brener et al14 who found differences of 6.6 cm for males and 6.9 cm for females. Shannon et al11 described tall males and females underestimating their weight more than their shorter counterparts but did not control for age.

**Weight**

Differences in mean self-reported and directly measured weights for the 9 studies that stratified by sex showed that females generally underestimated their weight more than their shorter counterparts, but did not control for age.

**Body Mass Index**

Two nationally representative studies6,8 and 3 convenience sample or locally based studies12,14,15 reported mean differences by sex for BMI (Table 2). Both males and females showed minimal mean differences in BMI (0 to 0.2 for males; −0.1 to −0.3 for females) in the nationally rep-

---

**Table 1. Characteristics of Studies in Reviewed Literature**

<table>
<thead>
<tr>
<th>Source</th>
<th>Sample Source</th>
<th>Sample Size (Males/Females)</th>
<th>Age Range, ya</th>
<th>Race/Ethnicity</th>
<th>Aware of DM/SR Data Before DM b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis and Gergen7</td>
<td>HHANES</td>
<td>829 (392/437)</td>
<td>12-19</td>
<td>Mexican American</td>
<td>Yes/Yes</td>
</tr>
<tr>
<td>Strauss4</td>
<td>NHANES III</td>
<td>1657 (767/890)</td>
<td>12-16</td>
<td>White/Hispanic, black, other</td>
<td>Yes/Yes</td>
</tr>
<tr>
<td>Himes and Faricya</td>
<td>NHANES III</td>
<td>1635 (759/876)</td>
<td>12-16</td>
<td>Non-Hispanic white, non-Hispanic black, Mexican American</td>
<td>Yes/Yes</td>
</tr>
<tr>
<td>Goodman et al8</td>
<td>Add Health</td>
<td>11 495 (5748/5747)</td>
<td>Grades 7-12</td>
<td>Non-Hispanic white, black, Hispanic, Asian, other</td>
<td>NR/No</td>
</tr>
<tr>
<td>Stewart9</td>
<td>6 Sites, Ohio, Washington, Massachusetts, South Carolina c</td>
<td>428 (NA)</td>
<td>14-17</td>
<td>85% White for ages 14-61 y</td>
<td>NR/Yes</td>
</tr>
<tr>
<td>Brooks-Gunn et al15</td>
<td>4 Private day schools, 1 city</td>
<td>151 (0/151)</td>
<td>11-13</td>
<td>White</td>
<td>Yes/Yes</td>
</tr>
<tr>
<td>Shannon et al11</td>
<td>School districts in Pennsylvania</td>
<td>726 (337/389)</td>
<td>11-13</td>
<td>White</td>
<td>NR/Yes</td>
</tr>
<tr>
<td>Himes and Story12</td>
<td>Indian reservation school in Minnesota</td>
<td>69 (41/28)</td>
<td>12-19</td>
<td>Native American</td>
<td>No/Yes</td>
</tr>
<tr>
<td>Hauck et al13</td>
<td>3 Indian Health Service areas</td>
<td>536 (NA/NA)</td>
<td>12-19</td>
<td>Native American</td>
<td>NR/Yes</td>
</tr>
<tr>
<td>Brener et al14</td>
<td>Subsample, 20 states and Washington, DC</td>
<td>2032 (957/1075)</td>
<td>Grades 9-12</td>
<td>White, black</td>
<td>Yes/Yes</td>
</tr>
<tr>
<td>Himes et al15</td>
<td>Comprehensive for 31 schools in Minnesota d</td>
<td>3797 (1936/1861)</td>
<td>12-18</td>
<td>White, black, Hispanic, Asian, other</td>
<td>NR/Yes</td>
</tr>
</tbody>
</table>

Abbreviations: Add Health, National Longitudinal Study of Adolescent Health; DM, directly measured; HANANES, Hispanic Health and Nutrition Examination Survey; NA, not available; NHANES, National Health and Nutrition Examination Survey; NR, not reported; SR, self-reported.

As reported in published articles. Includes highest year noted (eg, 12-16 includes up to 16.99 years).

Participants aware that they would/might be measured.

Rand Health Insurance Study.

Project EAT (Eating Among Teens).
representative studies, whereas there were greater underestimations (−1.2 to −2.3 for males; −1.0 to −3.0 for females) in the convenience sample or locally based studies.

Age or School Grade

Bias by age or school grade was reported by 4 studies. For the 3 studies that statistically examined bias by increasing age, 2 found increasing differences between self-reported and directly measured height,14,15 and 2 found decreasing differences between self-reported and directly measured weight,12,15 and 2 found decreasing differences between self-reported and directly measured BMI12,15; for each of these indexes, 1 of these 3 studies found no difference. Using a mixed-effects general linear model, Himes and Story12 reported that age was an independent predictor of reporting error for weight and BMI, but not for height, with the degree of underestimation of both self-reported weight and calculated BMI based on self-reported data decreasing with increasing age (errors less negative for each additional year in age by 0.78 kg for weight and 0.35 kg for BMI, respectively). Brener et al14 found that, after controlling for sex and race/ethnicity in linear regression analyses, each year’s increase in school grade was positively associated with an overestimate of self-reported height (β = 0.22; P < .001) and inversely associated with an underestimate of BMI based on self-reported values (β = −0.19; P = .002); they did not report any difference in weight bias by school grade.

Race/Ethnicity

Bias by self-reported race/ethnicity was examined statistically in 3 studies. Findings were disparate. Strauss1 found no statistically significant differences in reporting of height and weight by race/ethnicity (white/Hispanic individuals, black individuals, and individuals of other race/ethnicity) stratified by sex. After controlling for sex and school grade in their regression analysis, Brener et al14 found that white individuals overestimated their height more than black or Hispanic individuals or those of other racial or ethnic groups (β = 0.09; P < .001). Himes et al15 found significant effects for males only, with Asian individuals more accurately self-reporting height and underestimating weight more than white, African American, or Hispanic individuals or those of other racial or ethnic groups.

Weight Status

Six of the 11 studies examined reporting bias by weight status and found that overweight youth underestimated their weight7,11,13,15 and, if available, their resulting BMI7,12 more than normal-weight youth. For example, Strauss1 found that mean differences between self-reported and directly measured weight were significantly greater among overweight (BMI > 95th percentile) adolescents vs those not overweight (−4.6 kg vs 0.2 kg, respectively; P < .001). An additional study8 found underreporting of weight increased as directly measured weight increased for their entire sample (aged 14-61 years). Studies varied in these comparisons, making it impossible to examine bias by percentage of underestimation of directly measured weight by directly measured weight status.

DIFFERENCES IN PREVALENCE

We examined both absolute error (absolute difference between self-reported and directly measured prevalence) and relative error (absolute error divided by directly measured prevalence) in the 5 studies that had data available (Table 3). These studies showed that adolescent self-reported height and weight data underestimated prevalence of overweight, risk for overweight, or both, with wide variability in both absolute and relative er-

| Table 2. Sex-Specific Differences Between SR and DM Means for Height, Weight, and BMI |
|-----------------------------------------------|-----------------------------------------------|
| Differences in Means (SR−DM)                  |                                              |
| **Height, cm**                                | **Weight, kg**                                | **BMI**                                      |
| Source                                        | **M**                                        | **F**                                        | **M**                                        | **F**                                        |
| National Representative Surveys               |                                              |                                              |                                              |                                              |
| Strauss1                                      | −0.4                                         | −0.5                                         | 0.1                                          | −1.0                                         | NR                                           | NR                                           |
| Himes and Faricy6                             | −1.0                                         | −1.1                                         | −0.5                                         | −1.2                                         | 0.2                                          | NR                                           | 0.1                                          |
| Goodman et al12                               | NR                                           | NR                                           | −0.2                                         | −1.0                                         | 0                                            | 0                                            | 0.3                                          |
| Convenience Sample or Locally Based Surveys  |                                              |                                              |                                              |                                              |                                              |                                              |
| Brooks-Gunn et al19                          | NA                                           | 1.4                                          | NA                                           | −1.2                                         | NA                                           | NR                                           |
| Shannon et al11                              | 0.5                                          | −0.5                                         | −2.1                                         | −4.0                                         | NR                                           | NR                                           |
| Himes and Story12                            | 0.9                                          | 0.0                                          | −2.6                                         | −2.8                                         | −1.2                                         | −1.0                                         |
| Hauck et al15                                 | 1.5                                          | 0.5                                          | 1.5                                          | −2.1                                         | NR                                           | NR                                           |
| Brener et al14                               | 6.6                                          | 6.9                                          | −1.1                                         | −2.0                                         | −2.3                                         | −3.0                                         |
| Himes and Faricy18                           | 1.2                                          | 2.4                                          | −1.6                                         | −3.8                                         | −2.2                                         | −2.5                                         |
| All studies, range                            | −1.0 to 6.6                                   | −1.1 to 6.9                                   | −2.6 to 1.5                                   | −4.0 to −1.0                                   | −2.3 to 0.2                                   | −3.0 to −0.1                                   |

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); DM, directly measured; F, female; M, male; NA, not available, study included only females; NR, not reported; SR, self-reported.

a Stewart9 did not report if participants were aware of DM data or stratify by sex (height = 0.9 cm, weight = −1.1 kg).

b Means adjusted from mixed linear models by sex and included school, race, age, socioeconomic status, and categories of measured dimensions that were specific to each of the variables (eg, schools = random, stature = 6 categories).
rors. The wide variety of references and definitions of overweight used limited comparisons. For example, for the prevalence of overweight based on the Centers for Disease Control and Prevention 2000 growth charts, for the BMI cutoff of the 95th percentile or higher, absolute errors ranged from $-1.6\%$ to $-11.1\%$ and relative errors, from $-16.5\%$ to $-42.7\%$ (Table 3).

Table 4 shows sex-specific correlation coefficients between self-reported and directly measured height, weight, and BMI for 9 studies. The correlation coefficients for height, weight, and BMI were lower for females than males, with 3 exceptions. Including both males and females, correlation coefficients for height ranged from 0.62 to 0.91; for weight, 0.84 to 0.98; and for BMI, 0.79 to 0.93.

### CORRELATION COEFFICIENTS

Table 4 shows sex-specific correlation coefficients between self-reported and directly measured height, weight, and BMI for 9 studies. The correlation coefficients for height, weight, and BMI were lower for females than males, with 3 exceptions. Including both males and females, correlation coefficients for height ranged from 0.62 to 0.91; for weight, 0.84 to 0.98; and for BMI, 0.79 to 0.93.

### Age

Four studies examined self-reported and directly measured correlation coefficients by age, and in general found higher coefficients for the older adolescents, yet none reported statistical differences. For example,
Two nationally representative studies\(^7,8\) and 2 convenience sample studies\(^{13,14}\) examined the validity of using self-reported height and weight-calculated BMI to identify overweight (Table 5). The values for sensitivity among these 4 studies suggest that only between 55% and 75% of those adolescents who were truly overweight would be classified as overweight if their self-reported weight and height were used to determine BMI.

Three studies reported sensitivity by sex\(^{7,8,13}\) (Table 5), one reported by race/ethnicity\(^a\), and another reported by age\(^{13}\); no consistent patterns emerged for these subgroups. Using a BMI cutoff of the 95th percentile or higher, Goodman et al\(^8\) reported no difference in sensitivity by sex or race/ethnicity. Using a BMI cutoff of the 85th percentile or higher, Davis and Gergen\(^7\) reported no significant difference in sensitivity among Mexican American individuals by sex or by age (12-14 years, 15-17 years, or 18-19 years) for males and females (values not reported). Using a similar BMI cutoff for overweight, Hauck et al\(^13\) found that sensitivity for Native American individuals was lower among younger (12-14 years) than among older (15-19 years) males, yet higher among younger vs older females. Specificity was greater than 95% for the 4 studies, regardless of the cutoff or stratification (Table 5).

Table 5. Classification of Overweight by Self-reported vs Directly Measured Data: Sensitivity, Specificity, and Positive Predictive Value

<table>
<thead>
<tr>
<th>Source</th>
<th>Overweight Classification by BMI (Anthropometric References)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>National Representative Surveys</td>
<td>M</td>
<td>F</td>
<td>All</td>
</tr>
<tr>
<td>Davis and Gergen(^7)</td>
<td>≥ 85th Percentile (internal)</td>
<td>76.0</td>
<td>59.0</td>
<td>b,c</td>
</tr>
<tr>
<td>Goodman et al(^8)</td>
<td>≥ 95th Percentile(^{16,17})</td>
<td>74.0</td>
<td>70.2</td>
<td>a,c</td>
</tr>
<tr>
<td>Hauck et al(^13)</td>
<td>&gt; 85th Percentile(^{18})</td>
<td>76.2</td>
<td>70.4</td>
<td>d</td>
</tr>
<tr>
<td>Brener et al(^14)</td>
<td>≥ 85th Percentile(^{19})</td>
<td>NR</td>
<td>NR</td>
<td>60.5</td>
</tr>
<tr>
<td></td>
<td>≥ 95th Percentile(^{19})</td>
<td>NR</td>
<td>NR</td>
<td>54.9</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index; F, female; M, male; NR, not reported.
\(^a\)Not statistically different from males.
\(^b\)Reported no differences by age.
\(^c\)Reported no differences by race/ethnicity.
\(^d\)Did not conduct statistical testing of differences between males and females.

Our review identified 4 consistent patterns in bias across the studies when self-reported data were used to classify overweight among adolescents: (1) the prevalence of overweight was consistently underestimated by self-reported data, although the relative error ranged widely, from −2.4% to −42.7%; (2) sensitivity data showed that 25% to 45% of those overweight would be missed if self-reported data were used; (3) females underestimated their weight and BMI more than males; and (4) overweight youth underestimated their weight and BMI more than normal-weight youth.

We identified 3 additional issues, relevant to surveillance, targeting, program evaluation, and research, in need of further examination: (1) lack of or inconsistent information on other subgroup differences; (2) lack of information on whether biases are consistent over time; and (3) lack of consensus regarding an acceptable level of difference between self-reported and directly measured height and weight data and how this might vary depending on purpose of use.

The sensitivity data show that self-reported height and weight are relatively weak estimates of directly measured values for categorizing overweight status, which is similar or slightly worse than findings for adults. In comparison with adolescent data, the sensitivity data for adult US men and women aged 20 to 59 years are generally at the upper end of the range or higher than that for adolescents (Bowl et al\(^20\) 77% and 72%, respec-

Himes and Faricy\(^6\) reported variability by age stratified by sex and found lower values for 12-year-old boys and girls for height (0.57 vs 0.62, respectively) and BMI (0.70 vs 0.76, respectively) compared with 13- to 16-year-old boys and girls (both sexes, > 0.75 for height and > 0.83 for BMI), whereas Brener et al\(^14\) found high correlation coefficients (range, 0.86-0.96) for height, weight, and BMI for ninth- through 12th-grade adolescents.

**Race/Ethnicity**

Only Strauss\(^4\) and Brener et al\(^14\) examined correlation coefficients by race/ethnicity, stratified as white, black, and Hispanic individuals; Brener et al\(^1\) also included an “other” category. Strauss\(^4\) reported that BMI correlation coefficients for black individuals were significantly lower than those for white or Hispanic individuals (0.79 vs 0.89 and 0.85, respectively; \(P < .05\) for black vs white or black vs Hispanic individuals). Brener et al\(^14\) did not carry out significance testing but reported lower height correlations for Hispanic individuals than for individuals from the 3 other racial or ethnic groups (0.77 vs ≥ 0.88); BMI correlation coefficients for the “other” category were lower (0.79 vs ≥ 0.89).

**SENSITIVITY, SPECIFICITY, AND POSITIVE PREDICTIVE VALUE**

Two nationally representative studies\(^7,8\) and 2 convenience sample studies\(^{13,14}\) examined the validity of using self-reported height and weight-calculated BMI to identify overweight (Table 5). The values for sensitivity among these 4 studies suggest that only between 55% and 75% of those adolescents who were truly overweight would be classified as overweight if their self-reported weight and height were used to determine BMI.

Three studies reported sensitivity by sex\(^7,8,13\) (Table 5), one reported by race/ethnicity\(^a\), and another reported by age\(^{13}\); no consistent patterns emerged for these subgroups. Using a BMI cutoff of the 95th percentile or higher, Goodman et al\(^8\) reported no difference in sensitivity by sex or race/ethnicity. Using a BMI cutoff of the 85th percentile or higher, Davis and Gergen\(^7\) reported no significant difference in sensitivity among Mexican American individuals by sex or by age (12-14 years, 15-17 years, or 18-19 years) for males and females (values not reported). Using a similar BMI cutoff for overweight, Hauck et al\(^13\) found that sensitivity for Native American individuals was lower among younger (12-14 years) than among older (15-19 years) males, yet higher among younger vs older females. Specificity was greater than 95% for the 4 studies, regardless of the cutoff or stratification (Table 5).
tively; Kuczmarski et al,21 86%-95%; and Nieto-García et al,22 74%). Furthermore, the adult data show evidence of lower values among certain subgroups. For example, Bowlin et al20 found that for 20- to 29-year-old women, sensitivity was 60%, whereas Nieto-García et al22 found that for 60- to 69-year-olds, sensitivity was 64%.

In contrast to sensitivity analysis, we found that the correlation coefficients between self-reported and directly measured height, weight, and BMI were relatively high. However, correlation coefficients are strongly affected by extreme values, and they do not examine the accuracy of the classification of overweight status, our goal.

Additional research needs are to assess (1) the accuracy of self-reported data by order of self-reported and directly measured data collection; (2) whether advance knowledge of directly measured data influences self-reported data; (3) whether other potentially important factors not considered in this review, such as mental health status, socioeconomic status, or physical activity status, may also affect bias of self-reported data; and (4) comparisons of data based on standard accepted definitions of overweight status and anthropometric reference. One example of an additional important factor could be depression. overweight people are more likely to be depressed,23 and we found greater underestimates of weight among overweight adolescents. If depression is independently associated with underreporting of overweight status, then self-reported bias could also vary by level of depression.

Our findings raise questions about the practical implications for the use of self-reported data. The prevalence of overweight is assessed in populations for a variety of purposes, including identifying the problem, monitoring trends over time, program targeting, program evaluation, research, and advocating for funds for interventions. The utility of self-reported height and weight data depends on the purpose for which they are used. Self-reported data are almost always easier to collect and provide documentation of the problem of overweight even if the data are underestimates.

Prevalence based on self-reported data can be useful for monitoring overweight over time if the magnitude of biases remains constant over time. For short-term follow-up, the assumption of constant bias is probably reasonable, but no studies examined whether there have been changes in reporting bias over long periods. The fact that the degree of underreporting of prevalence varies widely among subgroups, from being almost undetectable in some studies (eg, males in Davis and Gergen7) to missing almost half of the problem in others,19 certainly raises the concern that reporting biases among subgroups are not constant across populations and settings.

The utility of self-reported data for targeting programs toward those at greatest risk depends on whether different population subgroups underreport BMI similarly. Females tend to underestimate their weight more than males, potentially leading to a false conclusion that males should be targeted more directly for overweight prevention programs. Heavier youth underestimate weight more than normal-weight youth, leading to an underestimate of the size of the target group. Other subgroup differences (in particular age, race/ethnicity, and missing data) have not been adequately studied, so it is impossible to know whether targeting based on self-reported data are appropriate.

Program evaluation of the effectiveness of obesity prevention and treatment interventions typically occurs over a relatively short time, such that it may be reasonable to expect that the degree of bias associated with self-reported data will not change between baseline and follow-up assessments. However, this expectation probably depends on the nature of the intervention, and one needs to think about the potential for bias based on the content of the intervention. For example, an intervention focused on BMI awareness or weekly monitoring of a person’s own weight would probably change that person’s knowledge of his or her directly measured weight and his or her ability to accurately report it. In such a case, evaluation might show increases in overweight because of the improved reporting even if the prevalence of overweight went down. Alternatively, there might be other factors associated with underreporting, such as the social desirability to be thin or poor body image, so that some participants would still underreport weight despite knowing their directly measured weight. Because the reasons for underreporting are multiple and complex, it is difficult to anticipate how an intervention might affect self-reported height and weight data.

Research often attempts to identify factors associated with overweight and nonoverweight as a dichotomous variable. To the extent that self-reported data misclassify adolescent overweight status and that this misclassification is unrelated to the factors of interest, research findings will be biased toward the null hypothesis. This bias implies that important factors associated with overweight will be overlooked. Where significant associations are found, however, one can expect that a real association does exist.

Where the estimated prevalence of overweight is high, the existence of self-reported height and weight data can be useful in advocating for programs and policies that address overweight and may be a prerequisite for securing funds to address the problem. However, policy makers and funders may be unimpressed with understated prevalence numbers and give inadequate attention to the problem, particularly in the long-term as the newness of the obesity epidemic passes.

One example of frequently used self-reported height and weight data is the Center for Disease Control and Prevention Youth Risk Behavior Surveillance System.24 This surveillance system regularly collects self-reported height and weight data and is often the only source of information on the prevalence of overweight among youth attending high school in a state. Without feasible alternative mechanisms to collect directly measured data, these self-reported data fill an important gap and are necessary for documenting the problem, advocating for funds for interventions, and monitoring trends over time.

Finally, this review has both strengths and limitations. To the best of our knowledge, this is the first literature review of this topic for adolescents. We identified 4 consistent biases and described the implications of these findings for using self-reported data. As for weaknesses, the lack of comparability among studies for cutoffs and references used to define overweight status, mea-
Researchers and public health practitioners should carefully consider the caveats described in this article when interpreting results based on self-reported data. Because of the limitations of relying on self-reported height and weight data for surveillance of overweight, public health practitioners should explore the feasibility of collecting directly measured heights and weights at the state and local levels for surveillance purposes. Historically, states collected directly measured height and weight data for a variety of purposes, particularly for screening. Arkansas measures the height and weight of children in schools for screening and surveillance. Missouri collected height and weight data during physical education classes as part of the curriculum for surveillance purposes. In New York, education law requires that children have health appraisals, including measurement of height and weight, at school entrance, as well as when students are in the second, fourth, seventh, and 10th grades. The New York Department of Health and some school districts are in discussions as to how to systematically collect and report these data to the Department of Health (Barbara A. Dennison, MD, written communication, April 2006). Experiences with these state programs, as well as other programs, related to data quality, representativeness, and cost could help determine the feasibility of developing state and local systems for collecting directly measured height and weight data for surveillance purposes.

Self-reported data are valuable if the only source of data. However, self-reported data underestimate overweight prevalence and there is bias by sex and weight status. Lower sensitivities of self-reported data indicate that one-fourth to one-half of those overweight would be missed. Other potential biases in self-reported data, such as differences across subgroups and inconsistency over time, need further clarification. The feasibility of collecting directly measured height and weight data on a state or community level should be explored because directly measured data are more accurate.

CONCLUSIONS

Accepted for Publication: May 24, 2007.
Correspondence: Bettylou Sherry, PhD, RD, Division of Nutrition and Physical Activity, Centers for Disease Control and Prevention, 4770 Buford Hwy, MS K25, Atlanta, GA 30341 (bsherry@cdc.gov).

Author Contributions: Study concept and design: Sherry, Jefferds, and Grummer-Strawn. Acquisition of data: Sherry and Jefferds. Analysis and interpretation of data: Sherry, Jefferds, and Grummer-Strawn. Drafting of the manuscript: Sherry and Jefferds. Critical revision of the manuscript for important intellectual content: Sherry, Jefferds, and Grummer-Strawn. Statistical analysis: Sherry and Grummer-Strawn.

REFERENCES