Received: 2 April 2021 | Revised: 15 May 2021 | Accepted: 20 May 2021

DOI: 10.1002/oby.23239

Obesity A Research Journal THE OBESITY SOCIETY WILEY

ORIGINAL ARTICLE

Epidemiology/Genetics

Evaluation of a suggested novel method to adjust BMI calculated from self-reported weight and height for measurement error

Katherine M. Flegal[1] | Barry I. Graubard[2] | John P. A. Ioannidis[1,3,4,5]

[1]Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, California, USA

[2]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA

[3]Department of Epidemiology and Population Health and Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, California, USA

[4]Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California, USA

[5]Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA

Correspondence
Katherine M. Flegal, Stanford Prevention Research Center, 1265 Welch Road, Stanford, CA 94305-5411, USA.
Email: kmflegal@stanford.edu, katherine.flegal@gmail.com

Funding information
METRICS has been funded from grants by the Laura and John Arnold Foundation.

Abstract

Objective: In 2019, Ward et al. proposed a method to adjust BMI calculated from self-reported weight and height for bias relative to measured data. They did not evaluate the adjusted values relative to measured BMI values for the same individuals.

Methods: A large data set (n = 37,439) with both measured and self-reported weight and height was randomly divided into two groups. The proposed method was used to adjust the BMI values in one group to the measured data from the other group. The adjusted values were then compared with the measured values for the same individuals.

Results: Before adjustment, 24.9% were incorrectly classified relative to measured BMI categories, including 7.9% in too high a category; after adjustment, 24.3% were incorrectly classified, with 12.8% in too high a category. The variance of the difference was unchanged. The adjustments reduced some errors and introduced new errors. At an individual level, results were unpredictable.

Conclusions: The suggested method has little effect on misclassification, can introduce new errors, and could magnify errors associated with factors, such as age, race, educational level, or other characteristics. State-level estimates and projections of obesity prevalence from values adjusted by this method may be incorrect.

INTRODUCTION

In 2019, Ward et al. (1) proposed a new method to correct for self-reporting bias in weight and height and to estimate state-specific and demographic subgroup–specific trends and projections of the prevalence of categories of BMI. Using their new approach, they published detailed state-specific estimates of obesity prevalence and projections of future levels by state in the United States, using data from the Behavioral Risk Factor Surveillance System (BRFSS). BRFSS is one of the sources (2) of nationally representative data among adults in the United States and it provides both national and state-specific estimates. However, BRFSS includes only self-reported weight and height data, which are well established to be inaccurate relative to measured data (3-5). The method of Ward et al. involved standardizing the overall distribution of BMI calculated from self-reported weight and height (called here “self-reported BMI”) from BRFSS values to the overall distribution of BMI values calculated from measured weight and height data (called here “measured BMI”) in the National Health and Nutrition Examination

Survey (NHANES). Both BRFSS and NHANES are nationally representative of the civilian noninstitutionalized US population (2,6-8), but unlike BRFSS, NHANES does not provide state-specific estimates.

Numerous comparison studies of self-reported versus measured weight and height show that self-reported weight and height have both random and systematic errors (3-5) relative to objectively measured weight and height. The systematic errors typically include greater underreporting of weight at higher measured weights and greater overreporting of height at lower measured heights. These errors can lead to what has been called the “flat slope syndrome,” a systematic tendency for high values to be underestimated and for low ones to be overestimated (9). This makes it difficult to predict the measured value from the self-reported value because a given self-reported value may arise either from a large measured value with high underreporting or a slightly smaller measured value with lower underreporting. This is similar to the problem identified by Plankey et al. (10).

The measurement errors in self-reported weight and height have also been found to have errors associated with other factors, including age, sex, race-ethnic group, region, income, education, recent physician visits, dieting behaviors, and health history (11-19), These errors generally may be affected by a host of issues, including social norms and social desirability biases. For example, data from several countries suggest that women with higher educational levels are more likely to underreport weight (17,19,20).

To standardize the distribution of self-reported BMI in BRFSS to the distribution of measured BMI in NHANES, Ward et al. calculated mean self-reported BMI from BRFSS by quantile of self-reported BMI and calculated mean measured BMI from NHANES by quantiles of measured BMI. They then calculated the difference in mean values between measured and self-reported BMI at each quantile and fit cubic splines to smoothly estimate self-report bias as a function of quantile across the entire BMI distribution. Each person’s BMI was then adjusted for this bias given his or her BMI quantile. The smoothed difference was added to the self-reported value in that quantile to produce an adjusted BMI value. They found that the resulting distribution of adjusted BMI in BRFSS was not significantly different from the distribution of measured BMI in NHANES according to the two-sample nonparametric Kolmogorov–Smirnov test.

Ward et al. describe their method as “adjustment for self-report bias.” They adjusted each person’s BMI. However, they did not demonstrate that the adjustments could provide correct values at the level of states, demographic subgroups, or individuals, because they did not compare measured BMI and adjusted BMI for the same individuals. Our objective was to investigate and describe some characteristics of the results of their method at the individual and subgroup level by applying the Ward method to a data set in which both measured and self-reported data are available for each individual.

Study importance

What is already known?

- BMI calculated from self-reported weight and height has both systematic and random error relative to measured data, but estimates of obesity prevalence based on measured weight and height are not available for individual US states.
- A proposed method to adjust for the error in self-report did not compare measured and adjusted BMI for the same individuals.

What does this study add?

- The proposed adjustment method had little or no effect on misclassification relative to measured BMI. It corrected some errors, introduced new errors, and could potentially magnify within- or between-state errors associated with factors such as age, race, educational level, or other characteristics.
- State-by-state comparisons and subgroup estimates of obesity prevalence and trends based on values adjusted by this method are not necessarily valid and should not be relied on.

How might these results change the direction of research or the focus of clinical practice?

- The limitations of BMI calculated from self-reported weight and height should be recognized.
- Methods are needed to adjust for the errors caused by using BMI calculated from self-reported weight and height.
- New methods should be validated against data for individual participants.

METHODS

In order to evaluate the 2019 method used by Ward et al., we used unweighted data from the NHANES surveys for survey cycles 1999 to 2000 through 2013 to 2014. These data sets are free and publicly available (21). Although these data come from a complex sample survey, here these data are used only to provide a large sample of self-reported and measured data for the same individuals. The difference between measured and self-reported BMI values was calculated as measured minus self-reported BMI; to reduce any effects of extreme errors in self-reported data, we trimmed the NHANES data by deleting values of the difference at or above the 99th percentile and at or below the 1st percentile, by sex and survey cycle. After trimming, there were 18,783 men and 18,656 women, ages 20 years and above, who had both measured and self-reported BMI values. We randomly divided the data set into two groups. We used the Ward et al. method to derive adjustment factors for self-reported BMI in Group 1 based on the distribution of measured BMI in Group 2. Because Group 1 and Group 2 are random samples from the same data set, this obviates the

TABLE 1 Descriptive information about the sample data set

	Group 1		Group 2	
	Men	Women	Men	Women
N	9,459	9,290	9,324	9,366
Mean (SD)				
Measured BMI	28.13 (5.56)	28.89 (7.00)	28.23 (5.58)	28.96 (7.18)
Self-reported BMI	27.75 (5.29)	28.06 (6.72)	27.83 (5.30)	28.12 (6.84)
Difference (measured minus self-reported)	0.37 (1.49)	0.84 (1.66)	0.40 (1.50)	0.83 (1.68)
Range (minimum, maximum)				
Measured BMI	14.14, 72.56	13.18, 71.30	14.86, 66.15	12.04, 82.95
Self-reported BMI	13.69, 66.41	13.52, 70.86	12.65, 61.27	12.21, 76.63
Difference (measured minus self-reported)	−4.79, 6.75	−5.04, 8.95	−4.73, 6.73	−4.97, 8.80
Correlations				
Measured BMI with self-reported BMI	0.946	0.960	0.947	0.962
Difference with measured BMI	0.368	0.326	0.375	0.352
Difference with self-reported BMI	0.102	0.089	0.112	0.121

need to have two different samples that are weighted to the same external population.

We applied the method described by Ward et al. to standardize the distribution of self-reported BMI in Group 1 to the distribution of the measured data in Group 2. The data in Group 1 were divided into sex-specific single percentile groupings, thus creating 100 groups for self-reported BMI. Within each percentile grouping of self-reported BMI, we calculated the sex-specific mean value of self-reported BMI. We also created 100 sex-specific percentile groups for measured BMI from the data in Group 2. Within each percentile grouping of measured BMI values, the sex-specific mean value of measured BMI was calculated. The mean difference between self-reported BMI from Group 1 and the measured BMI from Group 2 was calculated at the corresponding percentile of self-reported values. For example, the mean difference at the 50th percentile of self-reported values was calculated as the mean measured values for the 50th percentile of measured values in Group 2 minus the mean self-reported value for the 50th percentile of self-reported values in Group 1. The differences were smoothed over percentiles using cubic splines. The smoothed predicted mean difference for a given percentile category was added to the self-reported BMI values in that percentile category to create an "adjusted" BMI value for each individual within that category.

The adjusted values represent adjusted self-reported values within categories of self-reported values in Group 1. However, unlike the analysis of Ward, the data set that we used also provided the measured BMI values for the same individuals in Group 1. We compared the adjusted self-reported BMI values in Group 1 with the measured BMI values in Group 1 to evaluate the effects of the adjustments. Descriptive statistics, including means, differences, standard deviations, ranges, Pearson correlation coefficients, and Bland–Altman limits of agreement (LOA), were calculated. Data analysis and statistical procedures were carried out with SAS version 9.4 (SAS Institute, Cary, North Carolina). Group 2 was only used to calculate adjustment factors. All other analyses were carried out only on Group 1.

BMI was categorized according to the suggested groupings of <18.5 (underweight), 18.5 to 24.99 (normal weight), 25 to 29.99 (overweight), 30 to 34.99 (grade 1 obesity) and ≥35 (grades 2 to 3 obesity) (22). For both self-reported BMI and adjusted BMI, we estimated the proportion who were correctly classified relative to these measured BMI categories. Those who were misclassified were also categorized into two further groupings: those whose misclassified value put them in a higher BMI category than the measured value ("Too high") and those whose misclassified value put them in a lower BMI category ("Too low"). In almost all cases, the difference was only one BMI category; fewer than 1% of the sample had larger differences.

To provide subgroup examples, we divided the sample into groups based on sex, age group, race-ethnic group, income level, smoking status, disease history, alcohol consumption and educational level and various combinations of these factors. Within each subgroup, we calculated mean values for measured BMI, adjusted BMI, the difference between them (measured minus adjusted), obesity prevalence from measured BMI, and obesity prevalence from adjusted BMI.

For a sensitivity analysis, we repeated the analyses 5 more times using different random numbers each time to divide the data set into two groups.

RESULTS

Descriptive statistics for self-reported and measured BMI for both groups are shown in Table 1. The Kolmogorov–Smirnov test showed no significant difference between the distribution of measured BMI in the two groups ($p = 0.567$ for women and 0.554 for men). As expected, mean values of self-reported BMI were slightly lower than those for measured BMI. The difference between self-reported and measured BMI was correlated with measured values ($r = 0.371$ for men and 0.339 for women), showing the presence of systematic error, whereby the difference varies with the measured values. However, the correlation of the difference with self-reported values ($r = 0.107$ for men and 0.105 for women) was lower, likely because of the flat slope syndrome caused by systematic errors.

The adjustment procedure of Ward et al. is illustrated in Supporting Information Table S1 with examples of adjusting the data at the 25th, 50th, and 75th percentiles. In the 25th percentile category of self-reported BMI for men from Group 1, self-reported BMI values ranged from 24.06 to 24.17 with a mean of 24.12. In the 25th percentile category of measured BMI for men from Group 2, measured BMI values ranged from 24.25 to 24.40 with a mean value of 24.33, slightly higher than the mean self-reported value in the corresponding percentile of self-reported BMI. The difference between means for the 25th percentile is then 0.21 (24.33 minus 24.12); the smoothed difference was 0.25. As can be seen in the table, the difference between means is greater at the higher percentiles. The adjusted BMI values are created by adding the smoothed difference for the appropriate percentile to the self-reported BMI; the resulting difference between measured and adjusted values is close to zero.

The Group 1 data set contains measured and self-reported values for each individual, making it possible to assess some effects of the adjustment. In Supporting Information Table S2, the measured values of BMI in Group 1 within selected percentiles of self-reported BMI are displayed. Within a given percentile of self-reported BMI, the measured BMI varied over a range of roughly 3 to 4 BMI units lower than self-report to roughly 5 BMI units higher. The table displays the mean and standard deviations of the individual differences between the measured and self-reported BMI value and between the measured and adjusted BMI values. The mean differences were lower after the adjustments. However, the variance of the difference between measured and self-reported BMI was not changed by the adjustment.

The effects of adjustment on prevalence estimates and on misclassification relative to measured BMI are displayed in Table 2. Overall, almost a quarter of participants were misclassified. Relative to measured BMI categories, the self-reported BMI values tended to underestimate the prevalence of higher BMI categories and overestimate the prevalence of lower BMI categories. Relative to self-reported data, the adjusted prevalence estimates agreed much more closely with the measured prevalence values. The adjustment increased correct classification at the higher BMI levels but decreased correct classification in the normal weight category. Overall, the adjustment increased the proportion of people who were classified as too high, relative to measured BMI categories, and decreased the proportion who were classified as too low. A more detailed breakdown is shown in Supporting Information Table S3. Over 9% of men and 6% of women were misclassified into a higher category by self-reported BMI than by measured BMI; as shown in Supporting Information Table S3, the adjustment did not correct these misclassifications. The adjustment also created new misclassifications into a higher category as well, whereby some individuals who are correctly classified by self-reported BMI were misclassified upwards by adjusted BMI.

A specific example showing how these effects arise is presented in Supporting Information Table S4, showing data for the 60 men and women with a measured BMI in the range 29.00 to 29.04 (falling into the 63rd percentile of measured BMI for men or the 58th percentile of measured BMI for women). The corresponding values of self-reported BMI for these individuals spanned a wide range, from 24.4 to 31.7 for men and from 25.2 to 32.2 for women, equivalent to the 27th through 81st percentile of self-reported BMI for men and 39th through 78th percentile for women. The smoothed adjustment factor increased with the percentile of self-reported BMI. When the self-reported BMI percentile was low, the adjustment factor was

TABLE 2 Effects of adjustment of self-reported BMI on prevalence estimates and on misclassification relative to measured BMI by measured BMI category

		Measured BMI category				
		<18.5	**18.5 to 24.99**	**25 to 29.99**	**30 to 34.99**	**≥35**
Prevalence (%)						
Men	Measured	1.3	28.5	39.8	20.0	10.3
	Self-reported	1.1	31.0	40.1	18.8	9.0
	Adjusted	1.2	28.4	39.1	20.4	11.0
Women	Measured	2.1	30.9	29.5	20.1	17.4
	Self-reported	2.8	34.3	30.0	18.6	14.2
	Adjusted	2.3	30.9	29.5	19.3	18.1
Correct classification (%)						
Men	Unadjusted	44.1	82.4	75.0	66.2	76.9
	Adjusted	45.7	78.3	73.9	69.2	85.7
Women	Unadjusted	66.8	86.8	71.2	62.9	75.5
	Adjusted	59.2	82.5	71.7	65.9	87.0

small, and the adjusted value was still considerably below the true measured value. When the percentile of self-reported BMI was the same as the percentile of measured BMI, the adjustment factor was approximately correct, and the adjusted value was close to the measured value. When the percentile of self-reported BMI was higher than the percentile of measured BMI, the adjustment factor was too large, and the adjusted values were higher than the true measured values. As a result, in this case, after the adjustment, almost twice as many individuals were misclassified at the high end, with adjusted BMI values over 30. As these examples show, when the adjustment factors differ by percentile of self-reported data, then the effects of adjustment vary by self-reported value.

An example of the distortions from a different point of view is shown in Supporting Information Table S5, which contains data for all men and women with self-reported BMI values in the narrow range of 29.00 to 29.03. As the table illustrates, a narrow range of self-reported BMI may correspond to a wide range of measured BMI. All values of self-reported BMI were increased by the adjustment. When the measured values were lower than the self-reported values, the adjusted values were higher than the measured values. When the measured values were higher than the self-reported values, the adjusted values remained higher than the measured values for small differences but then lower than the measured values for larger differences. The adjustment factor was a fixed value for a given value of self-reported BMI. For men, who had a fixed adjustment factor of 0.52 in this narrow self-reported BMI range, all values of adjusted BMI were <30, regardless of the measured BMI value, leading to some classifications that were too low relative to measured BMI. In contrast, for women, who had fixed adjustment factors of 0.94 or 0.95, all values of adjusted BMI were >30, leading to some classifications that were too high relative to measured BMI.

The errors in self-reported data have been observed to vary by a variety of factors, including age, sex, race, region, educational level, and other factors. As an example, Table 3 shows the mean values of measured and adjusted BMI and of obesity prevalence according to measured or adjusted BMI for all combinations of sex, age group, smoking status, and race-ethnic group, for which there were at least 200 individuals. These and other characteristics differ widely across US states. As can be seen in Table 3, within each subgroup except for the oldest age, the mean adjusted values were higher than the mean measured values. In some cases, the absolute value of the difference was increased. In some cases, the prevalence of obesity was higher, sometimes as much as 4 percentage points higher, when adjusted values instead of measured values were used. In others, there was no difference in the prevalence of obesity, and in yet others, the prevalence of obesity was lower.

Summary results comparing the self-reported and the adjusted values are shown in Table 4. The overall misclassification rates were similar for self-reported and adjusted values. The proportion classified as “too low” decreased after the adjustment, and the proportion classified as “too high” increased. The mean difference between measured values and adjusted self-reported values was close to zero. However, the standard deviation of the difference was almost identical for the self-reported and adjusted values. Bland–Altman LOA show the range within which approximately 95% of the differences between self-reported and measured BMI would be expected to fall. The LOA for men for self-reported BMI were in the range of −3.18 to +3.86 BMI units and the limits for adjusted BMI were in the range of −3.75 to + .48 BMI units. The corresponding values for women were −2.99 to +4.63 for self-reported BMI and −3.82 to +3.88 for adjusted BMI.

As a sensitivity analysis, we repeated the analyses 5 times using different random numbers to divide the data sets into two different groups. Summary results are shown in Supporting Information Table S6 with results that are similar across the repetitions.

DISCUSSION

Here we describe some of the features of an adjusted estimate of self-reported BMI generated using the approach described by Ward et al. The values presented here are not intended as population estimates but simply to provide exploratory and descriptive information about the effects of such an adjustment method. These results should not be generalized to other data sets or to other BMI groupings. Misclassification rates will vary depending on the characteristics of self-reported BMI and also on the number of categories used. The characteristics of self-reported BMI vary widely across studies. As reported elsewhere, the distribution of self-reported BMI varies across national surveys of the same population (2).

Ward et al. adjusted self-reported data from BRFSS using measured data from NHANES for different individuals. This procedure relies on the assumption that the distribution of measured BMI data should be same in the two surveys because both are nationally representative (6-8). Instead of two different surveys, we divided the data set randomly into two groups. The distribution of measured BMI was not statistically different in the two groups, corresponding to the assumption required for the procedure by Ward et al. We used the procedure developed by Ward et al. to adjust the self-reported data in Group 1 to the measured data in Group 2. Because we had both measured and self-reported data for Group 1, we were able to compare the adjusted values for Group 1 with the measured values for the same individuals.

Self-reported BMI data commonly exhibit systematic errors, in which the error differs by the true value with greater measurement error at higher BMI values. As a result, when the self-reported value is too low relative to the measured value, the adjustment factor may be too low, and the adjusted BMI value may still be lower than the measured value. However, if the self-reported value is too high, the adjustment factor may be too high, and the adjusted BMI values may be higher than the measured value. Only if the self-reported BMI percentile is roughly the same as the measured BMI percentile will the adjustment factor be approximately correct. This is an inherent feature of the method when the difference increases with the measured value, an effect that is often observed (3-5,23-25).

TABLE 3 Effects of adjusted BMI within selected sex, age, smoking, and race-ethnic subgroups

Sex	Age group	Smoking	Race-ethnic group	Measured BMI	Adjusted BMI	Differences: Measured BMI minus self-reported BMI	Differences: Measured BMI minus adjusted BMI	Obesity prevalence by: Measured BMI	Obesity prevalence by: Adjusted BMI
Men	20 to 39	Never	MA	28.58	28.74	0.37	−0.16	31.64	34.63
	20 to 39	Never	NHW	28.16	28.32	0.35	−0.16	30.99	32.71
	20 to 39	Never	NHB	27.95	28.09	0.26	−0.15	31.13	31.13
	20 to 39	Current	NHW	26.32	26.43	0.21	−0.12	20.82	22.45
	20 to 39	Current	NHB	26.80	27.48	−0.20	−0.68	24.88	28.17
	40 to 64	Never	MA	29.13	29.40	0.29	−0.27	34.83	36.90
	40 to 64	Never	NHW	29.22	29.23	0.49	−0.02	35.58	35.44
	40 to 64	Never	NHB	29.94	30.08	0.40	−0.14	43.70	44.24
	40 to 64	Current	NHW	27.86	28.08	0.22	−0.22	29.81	32.09
	40 to 64	Current	NHB	26.71	27.51	−0.28	−0.80	27.12	29.49
	65+	Never	NHW	27.67	27.14	0.91	0.53	25.35	22.33
Women	20 to 39	Never	MA	28.09	28.16	0.74	−0.07	30.18	30.88
	20 to 39	Never	NHW	26.72	26.93	0.56	−0.21	24.96	24.66
	20 to 39	Never	NHB	30.59	30.69	0.96	−0.10	47.83	44.86
	20 to 39	Current	NHW	27.18	27.42	0.60	−0.24	30.43	30.19
	40 to 64	Never	MA	30.65	30.85	0.86	−0.20	50.32	51.38
	40 to 64	Never	NHW	28.90	29.06	0.72	−0.16	39.09	39.45
	40 to 64	Never	NHB	32.50	32.61	0.94	−0.11	56.67	56.87
	40 to 64	Current	NHW	27.78	28.02	0.65	−0.24	33.49	34.42
	40 to 64	Current	NHB	31.04	31.67	0.34	−0.63	51.24	54.73
	65+	Never	NHW	28.06	27.77	1.14	0.29	30.91	28.52
	65+	Never	NHB	30.35	30.27	1.07	0.09	48.10	48.10

Abbreviations: MA, Mexican American; NHB, Non-Hispanic Black; NHW, Non-Hispanic White.

TABLE 4 Summary effects of self-reported and adjusted BMI relative to measured BMI

	Correctly classified (%)[a]	Too low (%)	Too high (%)	Correlation of difference with measured values	Mean difference (SD)	Bland–Altman limits of agreement
Men						
Self-reported BMI	75.1	15.4	9.5	0.311	0.38 (1.81)	−3.16, 3.92
Adjusted BMI	75.1	11.4	13.6	0.159	−0.10 (1.84)	−3.70, 3.50
Women						
Self-reported BMI	75.0	18.6	6.4	0.278	0.84 (1.96)	−2.99, 4.63
Adjusted BMI	76.3	11.7	12.0	0.075	−0.06 (2.00)	−3.82, 3.88

[a] Percentages may not add to 100 because of rounding.

This adjustment does not reduce the variance of the difference. According to the Bland–Altman method of assessing agreement between two variables (26), the adjustments would change the numerical Bland–Altman LOA but would not narrow the distance between the upper and lower LOA. It should be noted that the LOA are high and indicate that regardless of the adjustment, a self-reported BMI would fall within a wide range of 6 to 7 BMI units around the measured value, suggesting relatively poor agreement on the level of the individual. After the adjustments, the adjusted BMI was higher than the measured BMI for 53% of men and 52% of women.

The method described by Ward et al. changes most values of self-reported BMI to a different adjusted value. At the overall population level, it produces approximately the same prevalence estimates as those from measured BMI. This result is predictable because their method is simply standardizing the distribution of one variable to the distribution of another variable. However, the adjustments do not necessarily produce correct values for individuals. The adjustments reduce some errors but also add new errors. As a result, the misclassification rates and the variance of the difference between self-reported and measured values show almost no change. As noted by Ward et al., their method preserves "the relative position of each person's BMI." (P. 2441). The method used by Ward et al. does not change the relative position of self-reported data. Therefore, it does not correct for any errors arising from factors such as age, education, race, health history, or any other variables. As a result, it is unlikely to correct accurately for state-specific and demographic subgroup errors. As a result of this and the inability of the NHANES to provide state-level BMI distributions to make state-specific corrections, it is unlikely that the method of Ward et al. can correct accurately for state-specific and demographic subgroup errors.

The method proposed by Ward et al. rests on the untested assumption that if the BMI distributions are standardized at the overall level, the adjusted results will therefore produce valid obesity prevalence estimates for subgroups. Their use of the adjusted values to rank, compare, and depict state-specific obesity prevalence rates rests on the assumption that the effects of adjustment are constant across states. The errors in self-reported weight and height have previously been shown to vary by demographic factors that also vary considerably across states. For example, the population proportion of people ages 65 and above ranges from over 20% in Maine and Florida to just over 11% in Alaska and Utah (27). The percent living in poverty ranges from over 20% in Mississippi and New Mexico to under 10% in Hawai'i, Maryland, and New Hampshire (28). In Maryland, the population is 29.4% non-Hispanic Black, whereas in Idaho, 0.6% of the population is non-Hispanic Black. The population of Maine is 93.4% non-Hispanic White, whereas in California, 37.0% of the population is non-Hispanic White (29). The prevalence of smoking ranges from 25% in West Virginia to 8.9% in Utah (30). If measurement errors differ by these and other factors, then the adjusted values may have errors that are different in different states. If one state, for example, has a preponderance of some demographic groups whose adjusted values tend to be biased upwards, and another state has a preponderance of groups whose adjusted values are biased downwards, then the estimates of obesity prevalence may be incorrect for both states, and comparisons between the two states may also be incorrect.

Measurement errors in exposure variables are an important issue (31-33). Unfortunately, the method suggested by Ward et al. does little to correct measurement errors and calls into question the accuracy of the detailed results reported in their article (1). Their method changes the overall mean values but accomplishes little else. Their method does not reduce misclassification, has little or no effect on the range of errors, and could potentially magnify errors associated with other factors such as age, race, educational level, health history, or other characteristics. At an individual level, almost every value is changed, but the results are unpredictable, and the degree of agreement is improved at the overall population level, not at the individual or subgroup level.O

CONFLICT OF INTEREST

The authors declared no conflict of interest.

ORCID

Katherine M. Flegal https://orcid.org/0000-0002-0838-469X

REFERENCES

1. Ward ZJ, Bleich SN, Cradock AL, et al. Projected U.S. state-level prevalence of adult obesity and severe obesity. *N Engl J Med*. 2019;381:2440-2450.
2. Flegal KM, Ogden CL, Fryar C, Afful J, Klein R, Huang DT. Comparisons of self-reported and measured height and weight, bmi, and obesity prevalence from National Surveys: 1999-2016. *Obesity (Silver Spring)*. 2019;27:1711-1719.

3. Connor Gorber S, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev*. 2007;8:307-326.
4. Maukonen M, Mannisto S, Tolonen H. A comparison of measured versus self-reported anthropometrics for assessing obesity in adults: a literature review. *Scand J Public Health*. 2018;46:565-579.
5. Seijo M, Minckas N, Cormick G, Comande D, Ciapponi A, Belizan JM. Comparison of self-reported and directly measured weight and height among women of reproductive age: a systematic review and meta-analysis. *Acta Obstet Gynecol Scand*. 2018;97:429-439.
6. Methodologic changes in the Behavioral Risk Factor Surveillance System in 2011 and potential effects on prevalence estimates. *MMWR Morb Mortal Wkly Rep*. 2012;61:410-413.
7. Chen TC, Parker JD, Clark J, Shin HC, Rammon JR, Burt VL. National Health and Nutrition Examination Survey: Estimation Procedures, 2011-2014. *Vital Health Stat*. 2018;2:1-26.
8. Mirel LB, Mohadjer LK, Dohrmann SM, et al. National Health and Nutrition Examination Survey: estimation procedures, 2007-2010. *Vital Health Stat*. 2013;2:1-17.
9. Kuskowska-Wolk A, Karlsson P, Stolt M, Rossner S. The predictive validity of body mass index based on self-reported weight and height. *Int J Obes (Lond)*. 1989;13:441-453.
10. Plankey MW, Stevens J, Flegal KM, Rust PF. Prediction equations do not eliminate systematic error in self-reported body mass index. *ObesRes*. 1997;5:308-314.
11. Cawley J, Maclean JC, Hammer M, Wintfeld N. Reporting error in weight and its implications for bias in economic models. *Econ Hum Biol*. 2015;19:27-44.
12. Craig BM, Adams AK. Accuracy of body mass index categories based on self-reported height and weight among women in the United States. *Matern Child Health J*. 2009;13:489-496.
13. Gillum RF, Sempos CT. Ethnic variation in validity of classification of overweight and obesity using self-reported weight and height in American women and men: the Third National Health and Nutrition Examination Survey. *Nutr J*. 2005;4:27.
14. Le A, Judd SE, Allison DB, et al. The geographic distribution of obesity in the US and the potential regional differences in misreporting of obesity. *Obesity (Silver Spring)*. 2014;22:300-306.
15. Villanueva EV. The validity of self-reported weight in US adults: a population based cross-sectional study. *BMC Public Health*. 2001;1:11.
16. Merrill RM, Richardson JS. Validity of self-reported height, weight, and body mass index: findings from the National Health and Nutrition Examination Survey, 2001-2006. *Prev Chronic Dis*. 2009;6:A121.
17. Jalkanen L, Tuomilehto J, Tanskanen A, Puska P. Accuracy of self-reported body weight compared to measured body weight. A population survey. *Scand J Soc Med*. 1987;15:191-198.
18. Masheb RM, Grilo CM. Accuracy of self-reported weight in patients with binge eating disorder. *Int J Eat Disord*. 2001;29:29-36.
19. Rossouw K, Senekal M, Stander I. The accuracy of self-reported weight by overweight and obese women in an outpatient setting. *Public Health Nutr*. 2001;4:19-26.
20. Pirie P, Jacobs D, Jeffery R, Hannan P. Distortion in self-reported height and weight data. *J Am Diet Assoc*. 1981;78:601-606.
21. National Center for Health Statistics. National Health and Nutrition Examination Survey. https://www.cdc.gov/nchs/nhanes/index.htm. Accessed April 2, 2021.
22. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: Executive Summary. Expert Panel on the Identification, Evaluation, and Treatment of Overweight in Adults. *Am J Clin Nutr*. 1998;68:899-917.
23. Stewart AW, Jackson RT, Ford MA, Beaglehole R. Underestimation of relative weight by use of self-reported height and weight. *Am J Epidemiol*. 1987;125:122-126.
24. Wright FL, Green J, Reeves G, Beral V, Cairns BJ. Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. *BMC Med Res Methodol*. 2015;15:81.
25. Palta M, Prineas RJ, Berman R, Hannan P. Comparison of self-reported and measured height and weight. *Am J Epidemiol*. 1982;115:223-230.
26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-310.
27. Himes CL, Kilduff L, Which US. Which states have the oldest populations? *Population Reference Bureau*. 2019. https://www.prb.org/which-us-states-are-the-oldest/ Accessed January 11, 2021.
28. Population Reference Bureau. Percent of the population living in poverty. https://www.prb.org/usdata/indicator/poverty/table/. Accessed January 11, 2021
29. State Population By Race, Ethnicity Data. https://www.governing.com/archive/state-minority-population-data-estimates.html. Accessed December 16, 2020.
30. World Population Review. Smoking Rates by State 2020. https://worldpopulationreview.com/state-rankings/smoking-rates-by-state. Accessed December 16, 2020.
31. Brakenhoff TB, Mitroiu M, Keogh RH, Moons KGM, Groenwold RHH, van Smeden M. Measurement error is often neglected in medical literature: a systematic review. *J Clin Epidemiol*. 2018;98:89-97.
32. Ranker LR, Petersen JM, Fox MP. Awareness of and potential for dependent error in the observational epidemiologic literature: a review. *Ann Epidemiol*. 2019;36:15-19.e12.
33. van Smeden M, Lash TL, Groenwold RHH. Reflection on modern methods: five myths about measurement error in epidemiological research. *Int J Epidemiol*. 2020;49:338-347.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Flegal KM, Graubard BI, Ioannidis JPA. Evaluation of a suggested novel method to adjust BMI calculated from self-reported weight and height for measurement error. *Obesity (Silver Spring)*. 2021;29:1700–1707. https://doi.org/10.1002/oby.23239