The ROYAL SOCIETY of MEDICINE

Research

Journal of the Royal Society of Medicine Open;
5(6) 1–7
DOI: 10.1177/2042533313514048

Comparison of bias resulting from two methods of self-reporting height and weight: a validation study

Melissa Scribani, Jessica Shelton, David Chapel, Nicole Krupa, Lynae Wyckoff and Paul Jenkins
Bassett Healthcare Network Research Institute, One Atwell Road, Cooperstown, NY 13326, USA
Corresponding author: Melissa Scribani. Email: melissa.scribani@bassett.org

Summary

Objectives: To contrast the validity of two modes of self-reported height and weight data.
Design: Subjects' self-reported height and weight by mailed survey without expectation of subsequent measurement. Subjects were later offered a physical exam, where they self-reported their height and weight again, just prior to measurement. Regression equations to predict actual from self-reported body mass index (BMI) were fitted for both sets of self-reported values. Residual analyses assessed bias resulting from application of each regression equation to the alternative mode of self-report. Analyses were stratified by gender.
Setting: Upstate New York.
Participants: Subjects ($n = 260$) with survey, pre-exam and measured BMI.
Main outcome measures: Prevalence of obesity based on two modes of self-report and also measured values. Bias resulting from misapplication of correction equations.
Results: Accurate prediction of measured BMI was possible for both self-report modes for men ($R^2 = 0.89$ survey, 0.85 pre-exam) and women ($R^2 = 0.92$ survey, 0.97 pre-exam). Underreporting of BMI was greater for survey than pre-exam but only significantly so in women. Obesity prevalence was significantly underestimated by 10.9% ($p < 0.001$) and 14.9% ($p < 0.001$) for men and 5.4% ($p = 0.007$) and 11.2% ($p < 0.001$) for women, for pre-exam and survey, respectively. Residual analyses showed that significant bias results when a regression model derived from one mode of self-report is used to correct BMI values estimated from the alternative mode.
Conclusions: Both modes significantly underestimated obesity prevalence. Underestimation of actual BMI is greater for survey than pre-exam self-report for both genders, indicating that equations adjusting for self-report bias must be matched to the self-report mode.

Keywords

research methods, epidemiology, public health, obesity

Introduction

Self-reported height and weight are commonly used in many areas of public heath research to calculate body mass index (BMI). Among the nationally known surveys that employ this method are the Behavioral Risk Factor Surveillance System (BRFSS)[1] and National Health Interview Survey (NHIS).[2] Because of its widespread use in discussions of the obesity epidemic, the validity of self-reported BMI is of great concern to public health researchers and policy-makers. More specifically, the magnitude of the obesity epidemic may be underestimated because of the use of self-reported BMI.

Although 'self-reported BMI' has been used as a general term to include any BMI derived from a subject's self-reported height and weight, there are actually a wide variety of ways in which these data can be obtained. The data may be provided on the phone, as in BRFSS[3] and the Canadian Community Health Survey,[4] by mail[5] or in a face-to-face interview, as in NHIS.[2]

The issue of how the self-reported height and weight were obtained may have particular relevance to the application of a correction factor. For example, a correction factor estimated from subjects who self-reported their height and weight in face-to-face interviews with full knowledge that they would subsequently be weighed and measured may not be appropriate to correct heights and weights obtained in a mail-in survey from subjects who did not expect to be measured.

Although the validity of BMI calculated from self-reported height and weight has been given considerable attention in the literature, there has been little systematic study of how the mode of self-report, and expectation of subsequent measurement, may affect validity. In one study comparing BRFSS self-reported and measured values, subjects were invited to participate in measurements only after the BRFSS interview was completed, precluding prior expectation that responses would be validated.[6] In the Canadian Community Health Survey,[4] which utilised computer-assisted telephone interviewing, it is not clearly stated whether the individuals had knowledge that their height and weight would subsequently be measured. This lack of clarification is common in much of the literature comparing self-reported to measured values.[3] Ezzati et al.,[3] discuss the importance of the mode of self-report on resulting bias but do not present results for self-reported data gathered by more than one mode on the same individual.

Data are presented here that contrast height and weight self-reported by mailed survey, wherein subjects had no expectation of subsequent measurement, and in face-to-face interview, after which subjects knew that their height and weight would be measured. It was hypothesised that self-report would yield underestimates of BMI in both cases, and that the underestimation would be significantly greater in cases where subjects had no expectation of subsequent validation of reported values.

Methods

The 2009 Upstate Health and Wellness Study comprised five substudies and covered the central New York counties of Otsego, Montgomery, Chenango, Schoharie, Delaware, Herkimer and Madison. These five substudies were Childhood Obesity, Adult Obesity, Access to Healthcare, Ageing and Household Health. Survey participants from the Ageing, Adult Obesity and Household Health studies were asked to self-report their height and weight on the survey without mention of subsequent exam. The analyses that follow relate to subjects from these three studies.

After completion of the five substudies, the researchers entered into a cooperative agreement with the National Kidney Foundation to recruit subjects for height and weight measurements that would be taken at the time of the kidney screenings. There were a total of eight kidney screening events that were geographically dispersed throughout the seven counties, subject to limitations posed by suitable clinic locations. Once the location of each screening was established, subjects who had self-reported their height and weight and who had a zip code within the immediate area were identified. A random sample of these subjects was sent a letter inviting them to attend the screening and receive a gift card for Walmart (worth $20 for individuals, $25 for families). A follow-up phone call was also used for recruitment.

After it was determined that the initial wave of attendees at the screening tended to be older, a more selective recruitment process targeting younger subjects was employed. To reach this younger group, a total of 10 additional limited screening events were held at clinics throughout the study region. These included measurement of weight, height and waist circumference only. Recruitment for this group was directed at respondents who were between the ages of 20 and 35 years. Initial advertisement for these screenings was conducted using email within the Bassett Healthcare Network. Further attempts at recruitment were made using telephone calls.

At these screening events, subjects were asked to self-report their height and weight with the knowledge that these would be subsequently measured during the exam. Following this self-report, each subject was weighed using a calibrated Tanita TBF 350 Pro Body Fat Composition Analyzer. Because this scale used an electronic current to gather percent body fat, participants with pacemakers were weighed using a regular digital scale. Height was measured using a calibrated stadiometer.

Statistical analyses

The design of the Household Health substudy was such that one person reported height and weight for all household members. Therefore, these analyses were limited only to those Household Health respondents who reported on themselves. Respondents to the Ageing and Adult Obesity surveys reported only on their own height and weight. Time between survey receipt and date of screening was recorded.

Subjects were classified as normal weight (18.5–24.99), overweight (25.0–29.99) and obese ($\geq$30.0) using both modes of self-report and also measured values. The distribution of measured BMI groups was contrasted with the distributions for both pre-exam and survey self-report. These distributions were compared using a series of gender-stratified McNemar's tests. The McNemar's tests were performed comparing survey versus measured values, pre-exam versus measured values and survey versus pre-exam values.

For each sex, multiple linear regression equations that included age were fit to estimate the subjects' actual BMI based on each mode of self-report. Analyses were also performed to demonstrate the bias that would result when the equation to predict measured BMI developed from one mode of self-report was used to correct the data obtained from the other mode. The resulting residuals (observed minus predicted values) were tested for a significant difference from zero, which is the null expectation for the mean of a set of residuals from a least squares regression.

Results

A total of 3820 individuals were invited to participate in the full scale screenings, with 423 (11.1%) participating. Among these 423 participants, 132 (31.2%) were excluded after it was determined that their height and weight on the survey had been reported by someone other than themselves. Subjects missing one or more of the main data points of interest (measured BMI, self-reported survey BMI or self-reported pre-exam BMI) were also excluded (4, 8 and 2 subjects, respectively). An additional 28 subjects were excluded because more than one year had elapsed between receipt of the survey and the date of the screening. Therefore, a total of 260 subjects were included in these analyses: 109 men and 151 women.

The mean age of men was 62.9 years (SD = 15.8), and the mean age of women was 59.5 (SD = 14.0) years. The mean measured BMI for both sexes was higher than either of the two self-reported values. For men, mean BMI was 27.8 (SD = 4.7), 28.1 (SD = 4.8) and 29.4 (SD = 5.0) for mailed survey, pre-exam and measured values, respectively. For women, these same values were 28.4 (SD = 7.2), 28.8 (SD = 7.2) and 29.8 (SD 7.8).

Results of the McNemar's tests obtained from Table 1 showed that the distribution of BMI group as ascertained by actual measurement differed significantly from the distribution obtained by both pre-exam (men $p < 0.001$, women $p = 0.007$) and survey self-report (men $p < 0.001$, women $p < 0.001$). Taken over all four cases (women pre-exam – mail, men pre-exam – mail), the results showed that in 88.2% of discordant pairs, the self-reported BMI category was lower than the measured category.

For the direct comparison of pre-exam versus survey self-report, approximately two-thirds of discordant pairs were those where the pre-exam BMI category was higher than the survey BMI category. This was true for both men (68.4%, $p = 0.11$) and women (65.5%, $p = 0.09$).

Table 1. Distribution of BMI according to three different assessment methods.

	Measured	Pre-exam self-report	Survey self-report
Men ($n = 109$)			
Normal weight	16.5%	24.4%	27.3%
Overweight	41.3%	44.3%	45.4%
Obese	42.2%	31.3%	27.3%
Women ($n = 151$)			
Normal weight	25.8%	26.5%	30.7%
Overweight	32.4%	37.0%	38.6%
Obese	41.9%	36.5%	30.7%

Table 1 shows the extent of underestimation that results from either of these two self-report modes. For men, the percent of obese individuals was underestimated by 10.9% and 14.9% by pre-exam self-report and survey self-report, respectively. For women, these same values were 5.4% and 11.2%.

When the survey self-reported values were used in the regression equations developed from the pre-exam values (Figures 1 and 3), the means of the residuals were significantly greater than zero for both genders using paired t-tests (men (Figure 1) = 0.404, $p = 0.013$, women (Figure 3) = 0.430, $p = 0.016$).

For men, the distribution of residuals contained an extreme outlier when the pre-exam self-reported values were used in the regression equation developed from the survey (Figure 2). Specifically, a value of + 15.03 was obtained, which was more than six standard deviations beyond the second highest value. Because of this, the Wilcoxon Signed Ranks Test was used to test the distribution of residuals. Despite this extreme positive outlier, both the mean (−0.191) and the median (−0.386) of these residuals were less than zero ($p = 0.004$ by Wilcoxon Signed Ranks Test). The distribution of residuals for women did not contain outliers (Figure 4). The mean value (−0.15) was again less than zero, but this difference did not reach significance by the paired t-test ($p = 0.173$).

Discussion

These data indicate that the mode of self-report (mailed survey vs. face-to-face interview) must be considered when correcting for self-report bias. Further, the subject's expectation of subsequent verification of his/her self-reported values is also important. Regardless of the mode of self-report, self-reported data on BMI that are not corrected for bias result in substantial underestimation of obesity levels.

The noteworthy strength of this study is that it is the only published report that directly compares two different modes of self-reporting BMI in the same subjects. The principal limitation was the low rate of participation (11.1%) which limits our ability to generalise to the study population. It is possible that those subjects who chose to participate in the kidney screening were more likely to have risk factors for kidney disease, including obesity, which may cause them to self-report height and weight differently from the general population.

Our investigation supports previous studies in finding significant underestimation of BMI via

Figure 1. Residual plot obtained when self-reported survey data are used to predict measured BMI in an equation derived from self-reported pre-exam data in men.

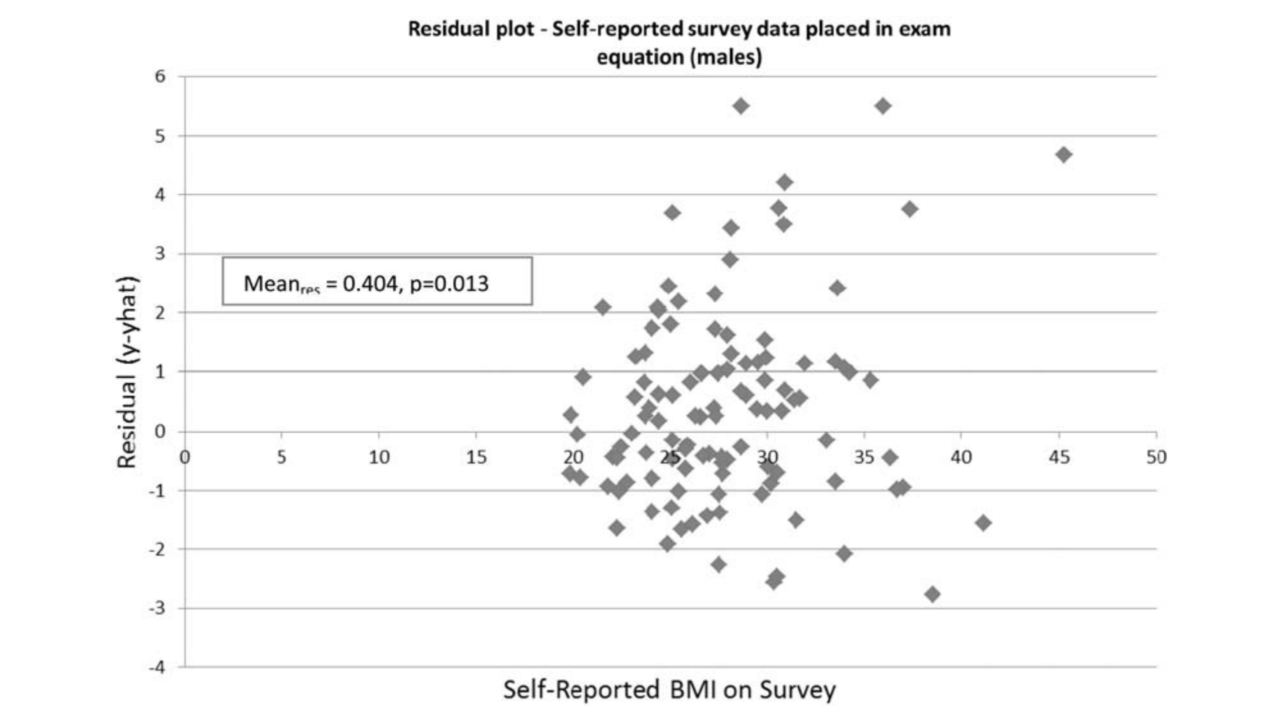


Figure 2. Residual plot obtained when self-reported pre-exam data are used to predict measured BMI in an equation derived from self-reported survey data in men.

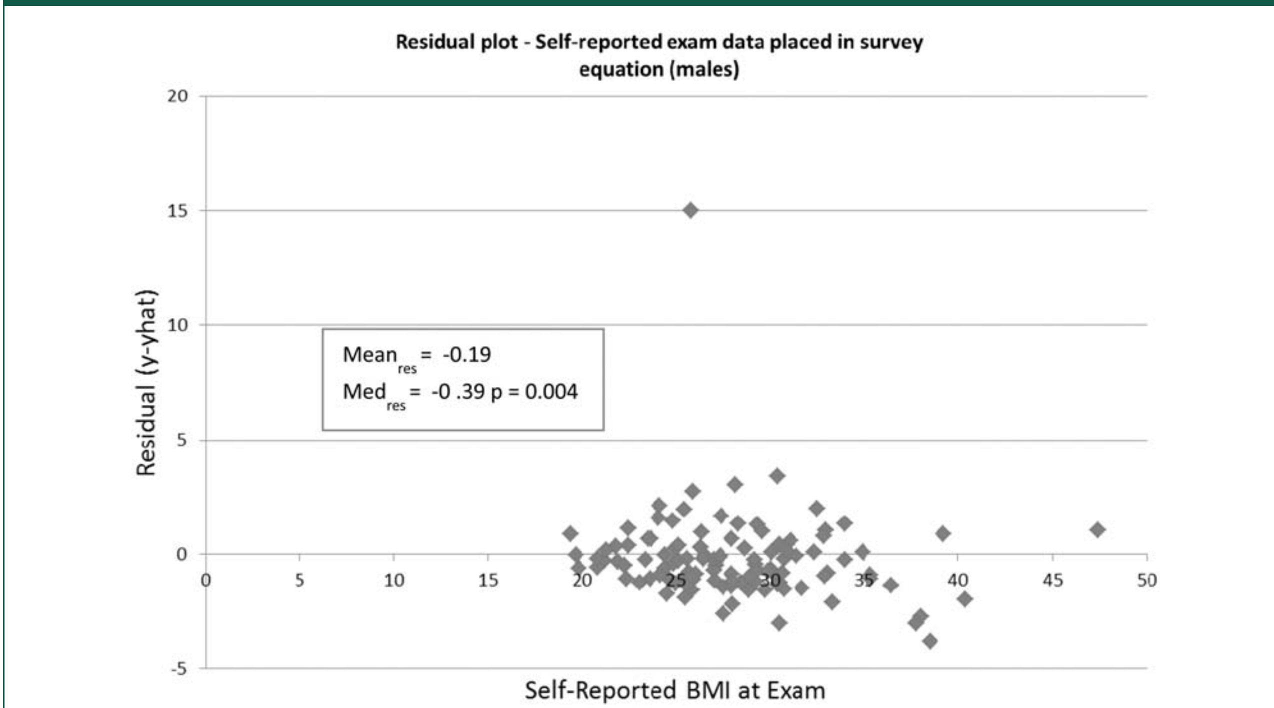


Figure 3. Residual plot obtained when self-reported survey data are used to predict measured BMI in an equation derived from self-reported pre-exam data in women.

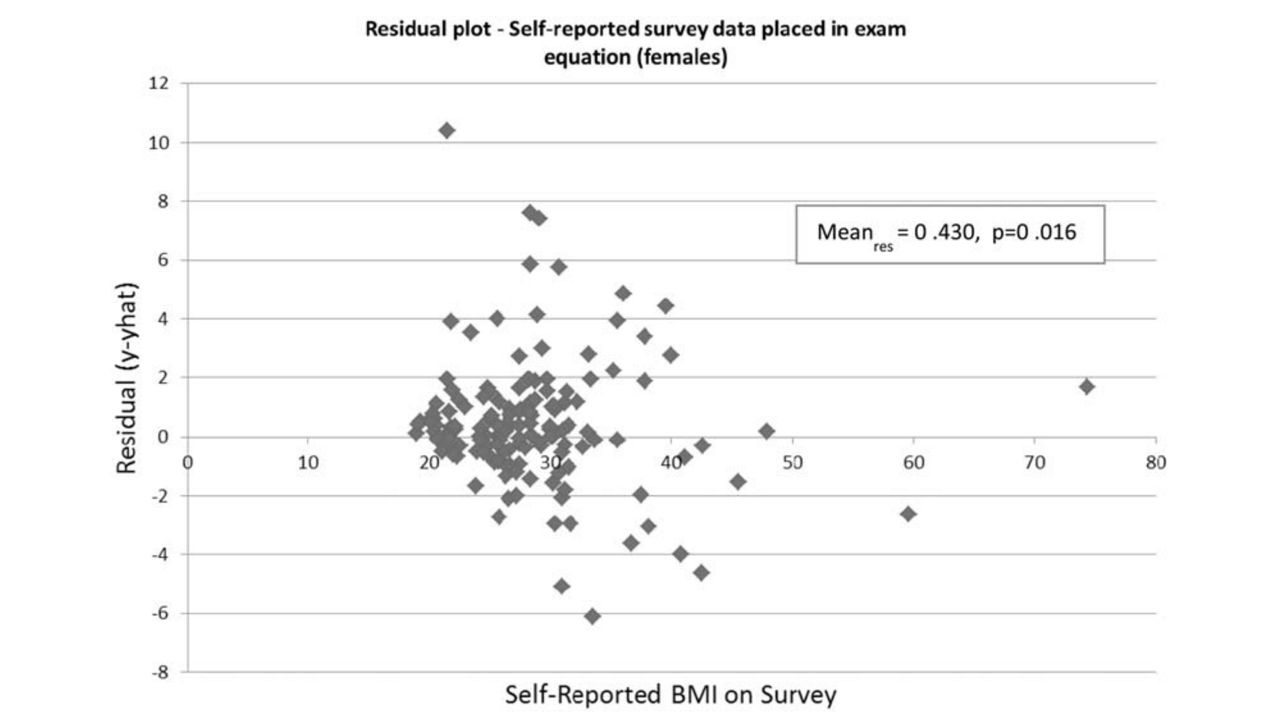
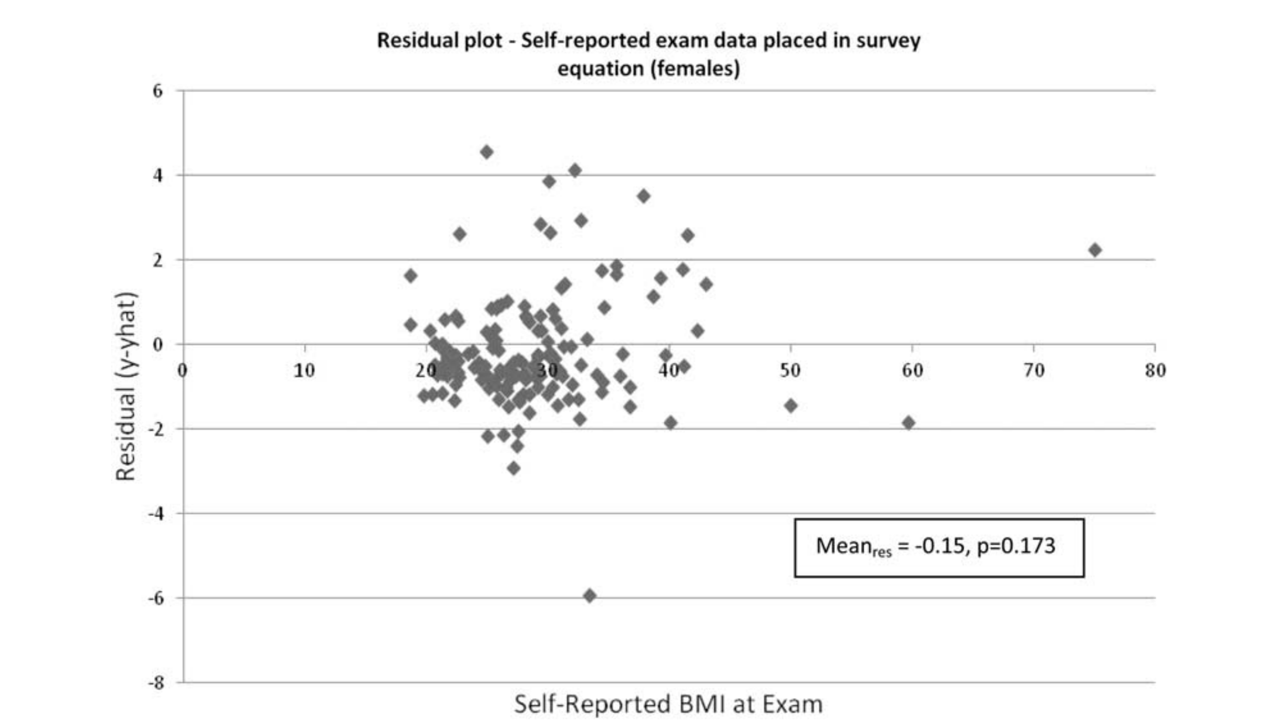


Figure 4. Residual plot obtained when self-reported pre-exam data are used to predict measured BMI in an equation derived from self-reported survey data in women.



self-report when compared to measured values.[3,4,7–9] Our results also verify previous findings that regression models can be developed to accurately predict BMI based on self-reported values.[3,4,8–10]

The expectation that height and weight values self-reported by mailed survey would yield BMI values significantly lower than those reported prior to anticipated measurement of height and weight was only partially borne out. Although the difference between survey and pre-exam self-report did not reach the traditional level of significance at $p = .05$, a trend towards greater underreporting of BMI in mailed survey self-report was seen in both men and women. These trends were considered to be of practical importance and to merit further investigation, as it must be acknowledged that the limited sample size of the study (151 women and 109 men) and consequently limited statistical power may be responsible for the failure of these results to reach statistical significance.

We propose two possible mechanisms for the trend towards greater underreporting of BMI on survey versus pre-exam self-report. If these two methods yield values that are truly not significantly different, it suggests that self-reported height and weight are simply inaccurate, regardless of the format in which it is obtained. This, therefore, implies that ignorance is the principal driver of loss of validity in these data. If, however, a more appropriately powered study demonstrates that BMI is indeed significantly underreported by survey versus pre-exam, this implies that deception is an important contributor to the loss of validity.

There were significant differences between the means of the residuals calculated by applying each correction formula to the self-reported data for which it was not specifically designed. Inputting survey self-report data into the model designed for pre-exam data generated positive residual values, indicating that the correction from the pre-exam regression was not large enough to accurately adjust the survey self-report data. Conversely, the survey self-report model generated a negative residual when used to correct pre-exam data, indicating that the correction was too large in this case, although not significantly so when applied to women alone.

We conclude that it is essential to correct BMI for self-report bias. When correcting these values, the estimating equations are most effective when they are specific to the mode of self-report. Future studies should explore the extent to which these equations may also benefit from taking other factors into account, such as education, ethnicity and socioeconomic status.

Declarations

Competing interests: None declared

Funding: This research was funded internally by the Bassett Healthcare Network Research Institute and received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval: All questionnaires, study procedures and protocols were approved by the Mary Imogene Bassett Hospital Institutional Review Board (Project #925, #927 and #929).

Guarantor: MS

Contributorship: MS assisted with data collection, conducted data analysis, drafted and revised the manuscript and managed the submission for publication; JS assisted with data analysis, drafted, revised and proofread the manuscript; DC drafted and revised the manuscript; NK assisted with data collection, managed all data, assisted with data analysis and drafted and revised the manuscript; LW coordinated all subject recruitment and data collection and revised the manuscript; PJ conceived the study, oversaw all data collection and analysis and supervised drafting and revising of the manuscript.

Acknowledgements: The authors wish to thank the National Kidney Foundation for their collaboration on this study.

Provenance: Not commissioned; peer-reviewed by Anwar Batieha

References

1. Anda R, Waller M, Wooten K, et al. Behavioral risk factor surveillance. *Morbidity and mortality weekly report (MMWR)* 1990; 39: 1–21.
2. Edwards WS, Winn DM, Kurlantzick V, et al. Evaluation of National Health Interview Survey Diagnostic Reporting. National Center for Health Statistics. *Vital Health Stat 2* 1994; 2: 11–18.
3. Ezzati M, Martin H, Skjold S, Hoorn SV and Murray CJL. Trends in national and state-level obesity in the USA after correction for self-report bias: analysis of health surveys. *J R Soc Med* 2006; 99: 250–257.
4. Akhtar-Danesh N, Dehghan M, Merchant AT and Rainey JA. Validity of self-reported height and weight for measuring prevalence of obesity. *Open Med* 2008; 2: 14–19.
5. Wada K, Tamakoshi K, Tsunekawa T, et al. Validity of self-reported height and weight in a Japanese workplace population. *Int J Obes* 2005; 29: 1093–1099.
6. Bowlin SJ, Morrill BD, Nafziger AN, Jenkins PL and Pearson TA. Validity of cardiovascular disease risk factors assessed by telephone survey: the Behavioral Risk Factor Survey. *J Clin Epidemiol* 1993; 46: 561–571.
7. McAdams MA, Van Dam RM and Hu FB. Comparison of self-reported and measured BMI as

correlates of disease markers in U.S. adults. *Obesity* 2007; 15: 188–196.
8. Nyholm M, Gullberg B, Merlo J, Lundqvist-Persson C, Rastam L and Lindblad U. The validity of obesity based on self-reported weight and height: implications for population studies. *Obesity* 2007; 15: 197–208.
9. Kuczmarski M, Kuczmarski R and Najjar M. Effects of age on validity of self-reported height, weight, and body mass index: findings from the third National Health and Nutrition Examination Survey, 1988–1994. *J Am Diet Assoc* 2001; 101: 28–34.
10. Jain RB. Regression models to predict corrected weight, height, and obesity prevalence from self-reported data: data from BRFSS 1999–2007. *Int J Obes* 2010; 34: 1655–1664.