Validity of child anthropometric measurements in the Special Supplemental Nutrition Program for Women, Infants and Children

Catherine M. Crespi, **Vivian H. Alfonso**, **Shannon E. Whaley**, and **May C. Wang**
Department of Biostatistics [C.M.C.], Department of Community Health Sciences [V.H.A., M.C.W.]
School of Public Health, University of California at Los Angeles, Los Angeles, CA 90024, USA.
PHFE WIC Program [S.E.W.], Irwindale, CA 91706, USA

Abstract

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) serves 50% of infants and 25% of preschool-aged children in the U.S. and collects height and weight measurements from eligible children every six months, making WIC data a valuable resource for studying childhood growth and obesity. We assessed the accuracy of measurements collected by WIC staff by comparing them to "gold standard" measurements collected by trained research staff. At seven WIC clinics in southern California, 287 children ages 2–5 years measured by WIC staff using WIC standard protocol were re-measured by research staff using a research protocol (duplicate measurements with shoes and outerwear removed taken by trained personnel). Intraclass correlation coefficients measuring agreement between WIC and research protocol measurements for height, weight and body mass index (BMI) were 0.96, 0.99 and 0.93, respectively. Although WIC measurements overestimated height by 0.6 cm and weight by 0.05 kg on average, BMI was underestimated by only 0.15 kg/m^2 on average. WIC BMI percentiles classified children as overweight/obese versus underweight/normal with 86% sensitivity and 92% specificity. We conclude that height, weight and BMI measurements of children aged 2–5 years collected by trained WIC staff are sufficiently accurate for monitoring and research purposes.

INTRODUCTION

Recent increases in the prevalence of childhood obesity in the United States and many other countries (1, 2) are a major public health concern. Childhood obesity tracks into adulthood (3, 4) and is associated with increased risk of psychological and respiratory conditions and with cardiovascular disease, cancer, diabetes and other chronic conditions later in life (5, 6). There is variation in child obesity rates by race/ethnicity, age, income and parental education, and the relationship of socio-demographic factors to child obesity is complex. For example, although obesity risk tends to be inversely associated with income and parental education among non-Hispanic white children, these associations are not observed consistently among Mexican American and non-Hispanic black children, aged 2–19 years, and may vary by gender (7). At preschool ages, obesity rates in the U.S. are highest among American Indians/Alaskan natives and Hispanics (8). Given the limited success of treatments (9) and the health consequences of childhood obesity, prevention in early life is key. Thus, there is a need for timely and reliable data sources for tracking growth in early childhood and identifying modifiable risk factors for childhood obesity.

Correspondence: Catherine M. Crespi, Ph.D., Department of Biostatistics, UCLA School of Public Health, 650 Charles E. Young Drive South, Room 51-254, Box 951772, Los Angeles, CA 90095-1772, Phone: 310-206-9364, Fax: 310-206-3566, ccrespi@ucla.edu.

In the U.S., the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) serves children up to age 5 years and pregnant, breastfeeding or postpartum women who are low income (<185% of the federal poverty level) and at nutritional risk (10). WIC serves nearly 50% of all infants and approximately 25% of all preschool-aged children in the United States (10). Regulations require that children who participate in WIC have their height and weight measured at least every 6 months. These measurements are used by WIC to assess program eligibility and to screen and intervene with families regarding potential health and nutritional risks such as excess weight.

The longitudinal anthropometric measurements of young children in the WIC program represent a large data resource for monitoring trends in childhood obesity and identifying risk factors. They have been used by the Centers for Disease Control and Prevention (CDC)'s Pediatric Nutrition Surveillance System (http://www.cdc.gov/pednss/what_is/pednss/index.htm) to monitor trends in child overweight, risk of overweight and other health indicators. WIC data have also been targeted for wider release as part of the Community Health Data Initiative of the U.S. Department of Health and Human Services and Institute of Medicine (http://www.hhs.gov/open/datasets/about.html), whose objective is to provide greater public access to health data. Given their nationwide coverage of an at-risk population and high concentrations of individuals in diverse geographic areas, anthropometric data in WIC administrative data systems also represent a largely untapped inexpensive resource for epidemiological research on childhood obesity, such as research on the impact of community food and physical activity environments (11, 12).

While the accuracy of height and weight measurements collected by WIC is critical to the success of current and potential future uses of WIC data for monitoring and research purposes, to our knowledge, the validity of these measurements has not been investigated. The objective of this study was to examine the accuracy of height, weight and body mass index (BMI) measurements derived from WIC administrative data systems for children ages 2 to 5 years. We sought to determine whether such measurements are sufficiently accurate to enable WIC administrative data systems to join other health data sources as valid and useful sources of information for tracking and investigating childhood obesity.

METHODS AND PROCEDURES

Study design

This study was the validation component of a larger project whose objective was to study the effects of neighborhood environment on the development of early childhood obesity using data from children aged 2–5 years extracted from PHFE WIC databases. PHFE WIC (http://www.phfewic.org) is the largest local WIC agency in the country and administers WIC clinics in Los Angeles, Orange and San Bernardino counties in Southern California. It serves over 300,000 clients every month, comprising 23% of California's and 4% of the nation's WIC participants. Over 80% of the families served by PHFE WIC are Hispanic.

The validation study sought to assess the accuracy of measurements obtained by WIC staff in accordance with their usual standard of care by comparing them to "gold standard" measurements obtained by research staff using standard measurement protocol. Accordingly, the study design was to have WIC staff measure and record the height and weight of children who were being recertified, have trained research staff obtain height and weight measurements of the same children using a standard research protocol, calculate BMI by both sets of measurements and compare the two sets of measurements. To ensure that the study was a validation of WIC's usual standard of care, WIC clinic staff were blinded to the true purpose of the study and not reminded of height/weight protocols nor retrained prior to

the study. WIC clinic staff were told that research staff were conducting a student-initiated survey that required taking height and weight measurements. The study protocol was approved by the Institutional Review Board of the University of California Los Angeles. Parents or guardians of participating children gave verbal consent before inclusion in the study.

Recruitment

Children were recruited from seven PHFE WIC sites in Los Angeles and Orange counties in California in the spring of 2010. WIC child participants must recertify for WIC eligibility every six months, at which time it is required that height and weight measurements be obtained by WIC staff. Children ages 1.9 to 5 years old who were attending the WIC clinic for the purpose of recertification were invited to participate in the study. The seven WIC sites were selected to achieve representation of the racial/ethnic composition of WIC clientele and a diversity of site characteristics such as staff turnover rates that might be associated with measurement accuracy.

Data collection

WIC staff—PHFE WIC paraprofessional clinic staff are trained upon hire to follow a written standardized height/weight measurement protocol and take required height and weight measurements of all children. Hands-on training with children is conducted in the clinic setting. To confirm adherence to height/weight protocols, staff are observed annually when taking height and weight measurements and asked to review the protocol if any deviations are observed. Seven to twelve staff members per site are involved in taking height and weight measurements.

WIC measurements—The WIC protocol for obtaining height and weight measurements requires WIC staff to either measure the child at the clinic during the recertification visit or obtain measurements from health care provider records (typically pediatrician visit records) brought to the recertification visit by the caregiver. The health care provider record can only be used if the visit was within 60 days of the WIC recertification appointment. WIC protocol calls for removal of shoes and outerwear before measuring children; height is measured to the nearest one-fourth inch and weight to the nearest one-fourth pound. At each of the WIC clinics in the study, WIC height measurements were obtained using a wall-mounted stadiometer (Model PE-WM-60-76, Prospective Enterprises, Portage, MI) and WIC weight measurements were obtained using a Health-O-Meter 402LB scale (Prospective Enterprises, Portage, MI). Scales are calibrated every 6 months by WIC administrative staff. Height and weight data obtained during recertification are entered into the California Integrated Statewide Information System (ISIS). For measurements obtained from health care provider records, the date of the measurement recorded in ISIS is the date of the provider visit. At PHFE WIC, about 80% of height and weight measurements are taken by WIC staff and about 20% are taken from pediatric provider records.

Research staff—Three research assistants, two of whom were bilingual, were trained to obtain height and weight measurements; another three were trained to obtain just weight measurements. The training involved use of a stadiometer (Shorr Board, Shorr Production, Olney, MD) for measuring height and a digital scale (Model HD-314, Tanita Corporation, Arlington Heights, IL) for measuring weight, using standardized protocols. Inter-observer error was assessed at the start of and midway through the 6-month data collection period, and was determined to be negligible (<0.1 cm for height and 0.1 kg for weight).

Research protocol measurements—During the child's recertification visit, research staff measured the child in a separate area of the WIC site, using the research study's

equipment as described above (Shorr Board and Tanita scales). The research protocol called for two measurements each of height and weight, taken by the same assistant; a third measurement was taken if the first two measurements differed by more than 0.5 cm (height) or 0.3 kg (weight). Standing height was measured to the nearest 0.1 cm, and weight was measured to the nearest 0.1 kg. The digital scale was calibrated once every 8 weeks using standard weights. The intraclass correlation coefficients (ICCs) for measurements within child were 0.9989 for heights and 0.9997 for weights, which suggests that intra-rater reliability was well above acceptable levels (13). Final height and weight values were obtained as means of the two or three measurements.

Statistical Analysis

Analysis was limited to children who had had WIC and research protocol measurements taken within 30 days of each other to minimize lack of concordance attributable to child growth. This resulted in the exclusion of some observations for which WIC measurements were based on pediatric provider records.

WIC height and weight measurements of children participating in this study were extracted from ISIS for purposes of comparison with measurements obtained by research staff. Child age, sex and race/ethnicity and preferred language of the family were also extracted from ISIS. For both WIC and research protocol measurements, height and weight measurements were used to compute BMI, defined as weight (kg)/height (m)^2, and height, weight and BMI percentiles were obtained using CDC sex- and age-specific reference values (14).

Statistical analyses assessed the reliability and accuracy of WIC measurements of height, height percentile, weight, weight percentile, BMI and BMI percentile as continuous variables, and the accuracy of WIC BMI percentiles to classify children as underweight or normal weight versus overweight or obese.

Several methods are available for assessing measurement reliability, with major methods including the ICC approach (15) and limits of agreement (LOA) (16). The first approach assesses reliability as the ICC, i.e., the percentage of the overall data variance that is due to true variability as opposed to variability due to measurement error or other sources. ICCs are unitless coefficients taking values between 0 and 1, which facilitates comparison across measurements on different scales. The LOA approach provides an interval within which 95% of differences between measurements by two raters are expected to lie. The limits are on the scale of the data, which facilitates interpretation of the clinical importance of differences but makes comparisons across different scales difficult. The two approaches provide complementary information and we provide both.

We calculated the 95% LOA for height, weight, BMI and their percentiles as mean difference $\pm$ 1.96 standard deviation of the differences (16), with differences calculated as WIC measurement minus research measurement. We estimated ICCs using variance components from the mixed model

$$Y_{ijk}=\mu+S_i+C_j+B_k+E_{ijk} \tag{1}$$

where μ is a global mean, S_i is a child random effect with mean 0 and variance σ_s^2, C_j is a clinic random effect with mean 0 and variance σ_c^2, B_k is a rater effect (WIC versus research protocol), and E_{ijk} is random error with mean 0 and variance σ_E^2. This approach is similar to Case 3 of Shrout and Fleiss (15) and Model C of Muller and Buttner (17), in which each target (i.e., child) is measured by a fixed number of raters (i.e., two raters, WIC and research

staff), but includes the addition of a clinic effect to account for sampling of children within a sample of clinics. Although research protocol measurements were taken by different research assistants, since inter-observer error among research staff was negligible as previously described, we did not include an additional term for research staff variance. Although B_k is a so-called fixed effect, it can be modeled as a random variable, with mean 0 and variance σ_B^2, and a consistent estimate of the ICC can be obtained as

$\sigma_S^2/(\sigma_S^2+\sigma_C^2+\sigma_B^2+\sigma_E^2)$(18, 19). We estimated ICCs for height, height percentile, weight, weight percentile, BMI and BMI percentile in the overall sample and within strata defined by age, sex, race/ethnicity, preferred language and source of WIC values (in-clinic versus provider record). We also estimated ICCs within strata defined by child height, weight, BMI and BMI percentile to assess whether reliability varied with child size. These strata were defined by dichotomizing the research protocol measurements approximately at their medians.

We assessed the accuracy of the WIC measurements as continuous variables by estimating the mean bias of the WIC measurements using the mixed model in equation (1) with the rater effect specified as fixed. We estimated mean bias for height, height percentile, weight, weight percentile, BMI and BMI percentile in the overall sample and within strata defined by age, sex, race/ethnicity, preferred language, source of WIC values, and child height, weight, BMI and BMI percentile. We tested for differences in mean bias among strata using likelihood ratio tests comparing models with and without an interaction between rater and factor.

We assessed the accuracy of WIC measurements as binary classifiers of overweight/obese versus underweight/normal by estimating sensitivity, specificity, positive and negative predictive value, and concordance between WIC and research protocol classifications. Underweight was defined as BMI below the 5th percentile, overweight as BMI ≥ 85th percentile and <95th percentile, and obesity as BMI ≥ 95th percentile (14). These estimates were obtained using mixed effects logistic regression models, with random intercepts for clinic. Models were formulated such that sensitivity, specificity, positive and negative predictive value and concordance could be estimated as marginal predictions. The delta method was used to obtain confidence intervals.

RESULTS

As shown in Figure 1, caregivers of 382 children who appeared to meet the age criteria were approached and consent was obtained for 367 (96%) children to participate in the study. Of these, 350 children met the age criteria (1.9–5 years). WIC measurements recorded within 30 days of measurement by research staff were obtained from 329 of these children, and 287 of these children cooperated with research staff to provide valid height and weight measurements (allowed their shoes and outerwear to be removed and provided at least two height and two weight measurements).

Table 1 summarizes characteristics of the children participating in the study. Boys and girls were about equally represented. About half of the children had parents who spoke Spanish as a preferred language, and 80% were Latino. These proportions are comparable to those for WIC families in California (20). BMI percentiles covered nearly the entire possible range, from 0.01 to 99.99.

Table 2 provides ICCs estimating the inter-rater reliability of WIC measurements in the sample overall and for strata defined by sex, race/ethnicity, preferred language (of parent), WIC site, and source of WIC measurements (WIC clinic or health care provider record), as well as by child height, weight, BMI and BMI percentile. Height, weight and BMI

measurements and percentiles were all highly reliable in the overall sample, with ICCs ranging from 0.90 to 0.99. In most strata, height measurements were somewhat less reliable than weight measurements, BMI measurements were less reliable than height or weight, and height, weight and BMI percentiles were less reliable than their corresponding raw values; African Americans and non-Hispanic whites were notable exceptions to these patterns. For almost all strata and measurements, ICCs were greater than 0.80, and most exceeded 0.90. Height measurements for African Americans and non-Hispanic whites and weight measurements for African Americans had lower reliability than for other race/ethnicities; this may have been a reflection of the clustering of these racial groups within certain clinics. In general, measurements from taller (≥100 cm) or heavier (≥16 kg) children were more reliable than measurements from shorter or lighter children; however, all ICCs in height and weight strata exceeded 0.80. The lowest ICCs were for BMI percentiles for children with BMIs ≥ 16.5 kg/m^2 or ≥75th percentile, which were in the 0.50–0.59 range. These ICCs may have been reduced due to attenuation of between-person variation within subgroups. The reliability of WIC measurements taken from provider records was lower than for in-clinic measurements especially for height-forage percentile (0.91 vs 0.63), and BMI (0.95 vs 0.64) and BMI percentile (0.91 vs 0.74).

Limits of agreement for the overall sample are presented in Table 3. About 95% of WIC height measurements were within ±2–3 cm of research protocol-measured height, 95% of WIC weights were within ±0.7–0.8 kg, and 95% of WIC BMIs were within about ±1 kg/m^2. The 95% LOA for height percentile was wider than that for weight percentile; 95% of WIC BMI percentiles were within about ±20 percentile points of research measurement-based BMI percentile.

Estimates of the mean bias of WIC measurements are also provided in Table 3. On average, WIC height measurements were 6 mm higher than research protocol measurements and WIC weight measurements were 54 grams higher. As a percentage of total height or total weight, the mean error in WIC measurements was 0.6% and 0.4%, respectively. WIC BMI measurements were 0.15 kg/m^2 lower than research protocol BMIs on average; this translated into a mean downward bias of 2.3 percentage points in BMI percentile. Bias was not significantly different by age, sex, race/ethnicity or preferred language, nor by child height or weight. Differences in bias were observed based on child BMI and BMI percentile. Children with higher BMI or BMI percentile tended to have their heights more overestimated than children with lower BMI or BMI percentile, but their weight percentiles were less overestimated; as a result, their BMI percentiles tended to have more downward bias. Some differences in bias were also detected by source of WIC measurements (weight and weight percentile).

Table 4 provides weight status classifications of the children (normal/underweight versus overweight/obese) based on research protocol measurements and WIC measurements. Only 5 children, i.e., 1.8% of the sample, met the criteria for underweight. Based on research protocol measurements, 39% (113/287) of the children were overweight or obese. Based on WIC measurements, 37% (105/287) would be classified as overweight or obese. Based on estimates from a mixed effects logistic regression model accounting for between-clinic variance, 86% (95% confidence interval (CI): 72%–99%) of overweight or obese children (by research protocol classification) were correctly classified as such by WIC measurements (sensitivity). About 92% (95% CI: 88%–97%) of normal or underweight children were correctly classified as such by WIC measurements (specificity). The positive predictive value of a WIC classification as overweight/obese was 87% (95% CI: 80%–93%), and the negative predictive value was 91% (95% CI: 83%–99%). About 89% (95% CI: 82%–95%) of the children were similarly classified by research protocol and WIC measurements (concordance).

DISCUSSION

Overall our findings suggest that height and weight measurements of children aged 2–5 years recorded by PHFE WIC staff and the BMIs and percentiles based on these measurements agree sufficiently well with "gold standard" research protocol measurements to support most research and monitoring purposes. ICCs for height, weight, BMI and corresponding percentiles were all 0.89 or higher in the overall sample, which is in the "almost perfect" range by the criteria of Landis and Koch (21) and the "substantial" reliability range by the criteria of Shrout (22). On average, WIC heights were only 6 mm higher than research protocol measurements and WIC weights were only 50 g higher. BMI percentile was underestimated by about 2.3 percentage points on average, and WIC BMI percentiles were less reliable for children with high BMIs (BMI ≥16 kg/m^2 or BMI percentile ≥75) than for children with low BMIs, with reliabilities of 0.50–0.59, which are in the "moderate" range by Landis and Koch (21) and the "fair" range by Shrout (22). The 95% limits of agreement for BMI percentile were fairly wide; however, the sensitivity and specificity of WIC measurements for classifying the children as overweight/obese were high, at 86% and 92% respectively, indicating that children were correctly classified as above or below the 85th BMI percentile with very good accuracy.

Of the measures we examined, BMI and BMI percentile are the quantities of greatest interest with regard to childhood obesity. BMI is ratio of weight to height squared, and its reliability is a function of the reliability of the height and weight measurements. Work of Cronbach (23) and others (24) has shown that the reliability of a ratio, as measured by the ICC, is typically lower than the reliability of the numerator and denominator, and this was evident in our overall sample ICCs. We found that WIC height errors were slightly greater than weight errors when expressed as a percentage of research protocol values. A multiplicative error in height measurement is more influential than an equivalent multiplicative error in weight in the calculation of BMI, which uses an inverse squared term for height. As a result, WIC BMIs, on average, were slightly underestimated. Overall, height measurements were somewhat less reliable than weight measurements, which could be due to the greater difficulty of obtaining compliance from young children during height measurements, which generally require longer periods of standing still than do weight measurements.

Insomuch as other WIC programs utilize similar height and weight protocols, our findings may be generalizable to other WIC programs. Informal assessment of protocols in New York, Illinois and Oklahoma suggest very similar protocols across regions, though in some cases measurements are taken to the 1/8 inch and 1/8 pound, which may increase measurement precision compared to the tolerances of 1/4 inch and 1/4 pound used in California.

Some of the upward systematic error in WIC height and weight measurements may be due to deviation from the WIC protocol of removing shoes and outerwear prior to measurement. We observed that implementation of the WIC standardized measurement protocol varied somewhat by site, which may have reduced the reliability of the measurements. As expected, accuracy and reliability tended to be lower for measurements derived from health care provider records compared to in-clinic WIC measurements. Provider measurements were typically taken days or weeks earlier than research protocol measurements; hence some discrepancy may be due to child growth, which would not reflect a lack of accuracy per se. However, some discrepancy may have reflected inaccuracy due to lack of use of a standardized protocol or calibrated equipment or inadequate training of health workers, which have been reported as concerns for child height measurement in health provider settings (25). Hence, source of measurements should also be considered when using anthropometric measurements from WIC.

It is informative to compare our findings with the accuracy of child height and weight measurements reported by other authors in other settings. A study in the United Kingdom that evaluated the accuracy of height and weight measurements of children at ages 4–43 months, taken from child health records, observed little systematic bias in these measurements but noted slight variations in accuracy between younger infants' and older children's measurements (14). We did not find differences by age in this group of children aged 2–5 years, but did find somewhat lower ICCs for shorter (<100 cm) and lighter (<16 kg) children. A study of the accuracy of height measurements of children taken in primary care clinics in the U.S. reported that 30% of measurements were within 0.5 cm of correct height (25). We found that about 36% (103/287) of WIC height measurements were within this tolerance.

Data for this study were collected from seven PHFE WIC sites in southern California and thus may not fully generalize to other WIC sites. This study did not examine the validity of height and weight measurements of infants aged 0 to 2 years nor of pregnant women; thus our findings should not be extended to infants and women participating in WIC. We conclude that height and weight measurements of young children aged 2–5 years taken at WIC sites using standardized protocols that include training and annual evaluations of staff adherence to measurement protocols are suitable for monitoring and research uses.

Acknowledgments

Abbreviations

BMI	Body mass index
CDC	Centers for Disease Control and Prevention
CI	Confidence interval
ICC	Intraclass correlation coefficient
ISIS	Integrated Statewide Information System
LOA	Limits of agreement
WIC	The Special Supplemental Nutrition Program for Women, Infants and Children

References

1. Wang Y, Lobstein T. Worldwide trends in childhood overweight and obesity. Int J Pediatr Obes. 2006; 1:11–25. [PubMed: 17902211]

2. Skelton JA, Cook SR, Auinger P, Klein JD, Barlow SE. Prevalence and trends of severe obesity among US children and adolescents. Acad Pediatr. 2009; 9:322–329. [PubMed: 19560993]

3. Reilly JJ, Methven E, McDowell ZC, Hacking B, Alexander D, Stewart L, Kelnar CJ. Health consequences of obesity. Arch Dis Child. 2003; 88:748–752. [PubMed: 12937090]

4. Guo SS, Wu W, Chumlea WC, Roche AF. Predicting overweight and obesity in adulthood from body mass index values in childhood and adolescence. Am J Clin Nutr. 2002; 76:653–658. [PubMed: 12198014]

5. Bjorge T, Engeland A, Tverdal A, Smith GD. Body mass index in adolescence in relation to cause-specific mortality: a follow-up of 230,000 Norwegian adolescents. Am J Epidemiol. 2008; 168:30–37. [PubMed: 18477652]

6. Reilly JJ, Kelly J. Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review. Int J Obes. 2011; 35:891–898.

7. Ogden CL, Lamb MM, Carroll MD, Flegal KM. Obesity and Socioeconomic Status in Children and Adolescents: United States, 2005–2008. NCHS Data Brief No 51. 2010

8. Sharma AJ, Grummer-Strawn LM, Dalenius K, Galuska D, Anandappa M, Borland E, Mackintosh H, Smith R. Obesity prevalence among low-income, preschool-aged children-United States, 1998–2008 (Reprinted from MMWR, vol 58, pg 769–773, 2009). JAMA. 2010; 303:28–30.

9. Kamath CC, Vickers KS, Ehrlich A, McGovern L, Johnson J, Singhal V, Paulo R, Hettinger A, Erwin PJ, Montori VM. Clinical review: behavioral interventions to prevent childhood obesity: a systematic review and metaanalyses of randomized trials. J Clin Endocrinol Metab. 2008; 93:4606–4615. [PubMed: 18782880]

10. Oliveira, V.; Frazao, E. The WIC Program: Background, Trends and Economic Issues. 2009. 2009.

11. Swinburn B. Obesity prevention in children and adolescents. Child Adolesc Psychiatr Clin N Am. 2009; 18:209–223. [PubMed: 19014868]

12. Sallis JF, Glanz K. Physical activity and food environments: solutions to the obesity epidemic. Milbank Q. 2009; 87:123–154. [PubMed: 19298418]

13. Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. Brit J Nutr. 1999; 82:165–177. [PubMed: 10655963]

14. Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z. 2000 CDC growth charts for the United States: methods and development. Vital Health Stat. 2002; 11:1–190.

15. Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979; 86:420–428. [PubMed: 18839484]

16. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1:307–310. [PubMed: 2868172]

17. Muller R, Buttner P. A critical discussion of intraclass correlation coefficients. Stat Med. 1994; 13:2465–2476. [PubMed: 7701147]

18. Carrasco J, Jover L. Estimating the generalized concordance correlation coefficient through variance components. Biometrics. 2003; 59:849–858. [PubMed: 14969463]

19. Rousson V. Assessing inter-rater reliability when the raters are fixed: two concepts and two estimates. Biom J. 2011:2011.

20. PHFE WIC. [(Accessed June 30, 2011)] PHFE WIC Program. 2010. Available at: http://www.phfewic.org/PHFEWIC.htm

21. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

22. Shrout P. Measurement reliability and agreement in psychiatry. Stat Methods Med Res. 1998; 7:301–317. [PubMed: 9803527]

23. Cronbach LJ. The reliability of ratio scores. Ed Psych Meas. 1941; 1:269–278.

24. Allison DB, Paultre F, Goran MI, Poehlman ET, Heymsfield SB. Statistical considerations regarding the use of ratios to adjust data. Int J Obes (Lond). 1995; 19:644–652.

25. Lipman TH, Hench KD, Benyi T, Deluane J, Gilluly KA, Johnson L, Johnson MG, McKnight-Menci H, Shorkey D, Shults J, Waite FL, Weber C. A multicentre randomised controlled trial of an intervention to improve the accuracy of linear growth measurement. Arch Dis Child. 2004; 89:342–346. [PubMed: 15033843]
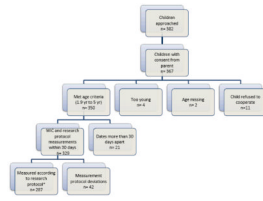
Figure 1.
Flowchart of participant inclusion. *Refers to children who allowed their shoes and outerwear to be removed and provided at least two height and two weight measurements to research staff.

Table 1

Characteristics of children participating in the study (N = 287)

Characteristic	N	%
Gender		
Female	136	47
Male	151	53
Ethnicity		
African American	24	8
Asian	7	2
Hispanic	230	80
White	26	9
Preferred Language		
English	146	51
Spanish	141	49
Age		
1.9–2.9 years	87	30
3.0–3.9 years	104	36
4.0–4.9 years	96	33
	Median	Range
Height (cm)	98.6	81.0 – 117.7
Weight (kg)	16.1	10.4 – 29.0
BMI (kg/m^2)	16.6	12.5 – 26.6
BMI percentile	76.4	0.01 – 99.99

Table 2

Intraclass correlation coefficients assessing inter-rater reliability of WIC usual care measurements, overall and by child factors and source

Factor	N (%)	Height (cm)	Height for Age Percentile	Weight (kg)	Weight for Age Percentile	BMI (kg/m²)	BMI Percentile
Overall (with 95% confidence intervals)	287 (100%)	.96 (.92, 1.00)	.89 (.82, .95)	.99*	.97 (.94, 1.00)	.93 (.92, .95)	.90 (.88, .92)
Age of child							
2–3 years	87 (30%)	.90	.83	.98	.94	.93	.84
3–4 years	103 (36%)	.92	.89	.99	.98	.94	.89
4–5 years	97 (34%)	.94	.95	.96	.96	.90	.86
Sex of child							
Female	136 (47%)	.96	.91	.99	.98	.93	.93
Male	151 (53%)	.97	.86	.99	.96	.93	.87
Race/ethnicity of child							
African American	23 (8%)	.77	.92	.55	.99	.93	.95
Asian **	7 (2%)	--	--	--	--	--	--
Hispanic	230 (80%)	.97	.89	.99	.98	.94	.90
Non-Hispanic white	27 (9%)	.71	.81	.88	.97	.80	.85
Preferred language of family							
English	146 (51%)	.97	.92	.99	.97	.93	.88
Spanish	141 (49%)	.95	.89	.98	.98	.91	.91
Child height§							
<100 cm	167 (58%)	.89	.86	.98	.97	.89	.87
≥100 cm	120 (42%)	.92	.92	.98	.99	.95	.93
Child weight§							
<16 kg	137 (48%)	.91	.85	.92	.94	.82	.81
≥16 kg	150 (52%)	.93	.88	.98	.97	.93	.98
Child BMI§							
<16.5 kg/m²	126 (44%)	.99	.96	.99	.97	.86	.86
≥16.5 kg/m²	161 (56%)	.92	.83	.99	.95	.88	.59
Child BMI percentile§							
<75th	140 (49%)	.98	.92	.98	.97	.83	.81

Factor	N (%)	Height (cm)	Height for Age Percentile	Weight (kg)	Weight for Age Percentile	BMI (kg/m^2)	BMI Percentile
≥75th	147 (51%)	.91	.87	.99	.95	.89	.50
Source of WIC values							
In-clinic	268 (93%)	.97	.91	.99	.98	.95	.91
Provider record	19 (7%)	.89	.63	.94	.95	.64	.74

*
Confidence interval could not be computed due to extremely high agreement.

**
Not reported due to small sample size

§
The median was used to dichotomize these variables

Table 3

Limits of agreement of research protocol and WIC usual care measurements and estimated mean bias of WIC usual care measurements, overall and by child factors and source

Factor	N (%)	Height (cm)	Height for Age Percentile	Weight (kg)	Weight for Age Percentile	BMI (kg/m^2)	BMI Percentile
				95% limits of agreement			
Overall	287 (100%)	(−2.1, 3.3)	(−15, 23)	(−0.7, 0.8)	(−9, 11)	(−1.3, 1.0)	(−21, 18)
				Estimated mean bias			
Overall (with 95% confidence intervals)	287 (100%)	0.6 (0.4,0.8)	4.0 (2.7, 5.2)	0.05 (−0.01,0.11)	0.8 (0.2,1.5)	−0.15 (−0.23,−0.07)	−2.3 (−3.6,−0.9)
Age of child							
2–3 years	87 (30%)	0.4	3.5	0.02	0.6	−0.1	−2.6
3–4 years	103 (36%)	0.7	4.8	0.04	0.6	−0.2	−3.7
4–5 years	97 (34%)	0.7	3.5	0.10	1.3	−0.1	−0.4
P-value		*NS*	*NS*	*NS*	*NS*	*NS*	*NS*
Sex of child							
Female	136 (47%)	0.5	3.1	0.04	1.1	−0.1	−1.2
Male	151 (53%)	0.7	4.7	0.07	0.6	−0.2	−3.2
P-value		*NS*	*NS*	*NS*	*NS*	*NS*	*NS*
Race/ethnicity of child							
African American	23 (8%)	1.0	5.3	0.07	0.4	−0.3	−2.4
Asian*	7 (2%)	--	--	--	--	--	--
Hispanic	230 (80%)	0.6	3.8	0.07	1.0	−0.1	−2.0
Non-Hispanic white	27 (9%)	0.6	3.2	−0.09	− 0.8	−0.3	−4.1
P-value		*NS*	*NS*	*NS*	*NS*	*NS*	*NS*
Preferred language of family							
English	146 (51%)	0.6	3.8	0.04	0.2	−0.2	−2.7
Spanish	141 (49%)	0.5	4.1	0.07	1.5	−0.1	−1.8

Factor	N (%)	Height (cm)	Height for Age Percentile	Weight (kg)	Weight for Age Percentile	BMI (kg/m²)	BMI Percentile
P-value		*NS*	*NS*	*NS*	*NS*	*NS*	*NS*
Child height							
<100 cm	167 (58%)	0.6	4.8	0.07	1.2	−0.1	−2.5
≥100 cm	120 (42%)	0.6	2.7	0.03	0.3	−0.2	−1.9
P-value		*NS*	*NS*	*NS*	*NS*	*NS*	*NS*
Child weight							
<16 kg	137 (48%)	0.5	4.6	0.08	1.3	−0.1	−2.4
≥16 kg	150 (52%)	0.7	3.4	0.03	0.4	−0.2	−2.2
P-value		*NS*	*NS*	*NS*	*NS*	*NS*	*NS*
Child BMI							
<16.5 kg/m²	126 (44%)	0.3	2.2	0.08	1.6	0.01	−0.4
≥16.5 kg/m²	161 (56%)	0.9	5.3	0.03	0.2	−0.28	−4.0
P-value	*NS*	*< .05*	*< .05*	*NS*	*< .05*	*<.001*	*< .05*
Child BMI percentile							
<75th	140 (49%)	0.4	3.2	0.10	1.69	−0.02	−1.0
≥75th	147 (51%)	0.8	4.7	0.01	0.04	−0.28	−3.5
P-value	*NS*	*< .05*	*NS*	*NS*	*< .05*	*< .05*	*NS*
Source of WIC values							
In-clinic	268 (93%)	0.6	4.1	0.07	1.1	−0.1	−1.9
Provider record	19 (7%)	0.4	2.4	−0.23	−2.1	−0.4	−7.1
P-value	*NS*	*NS*	*NS*	*< .05*	*< .05*	*NS*	*NS*

Positive mean bias values indicate that WIC measurements were overestimates on average; negative mean bias values indicate that WIC measurements were underestimates on average. P-values obtained from likelihood ratio tests comparing models with and without an interaction between rater and factor.

NS: difference is not statistically significant

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Comparison of weight status classifications defined using research protocol measurements and WIC usual care measurements of 287 children

		Classification based on WIC usual care measurements		Total
		Underweight or normal weight	Overweight or obese	
Classification based on research protocol measurements	Underweight or normal weight	160	14	174 (61%)
	Overweight or obese	22	91	113 (39%)
	Total	182 (63%)	105 (37%)	287 (100%)