Reproducibility and Intermethod Reliability of a Calcium Food Frequency Questionnaire for Use in Hispanic, Non-Hispanic Black, and Non-Hispanic White Youth

Nicholas J. Ollberding, PhD; Vicente Gilsanz, MD; Joan M. Lappe, PhD; Sharon E. Oberfield, MD; John A. Shepherd, PhD; Karen K. Winer, MD; Babette S. Zemel, PhD; Heidi J. Kalkwarf, PhD, RD

ARTICLE INFORMATION

Article history:
Accepted 23 December 2014
Available online 12 February 2015

Keywords:
Diet
Questionnaires
Calibration
Diet recall
Nutrition assessment

Supplementary materials:
The Figure and Table 5 are available online at www.andjrnl.org

http://dx.doi.org/10.1016/j.jand.2014.12.016

ABSTRACT

Background A dietary assessment instrument designed for use in a nationally representative pediatric population was required to examine associations between calcium intake and bone mineral accrual in a large, multicenter study.

Objective To determine the reproducibility and intermethod reliability of a youth calcium food frequency questionnaire (FFQ) in a multiracial/ethnic sample of children and adolescents.

Design Reproducibility (n=69) and intermethod reliability (n=393) studies were conducted by administering repeat FFQs and three unannounced 24-hour dietary recalls to stratified random samples of individuals participating in the Bone Mineral Density in Childhood Study.

Participants/setting Children and adolescents ages 5 to 21 years.

Main outcome measures Calcium intake estimated from the FFQ and 24-hour dietary recalls.

Statistical analysis Reproducibility was assessed by the intraclass correlation coefficient (ICC). Intermethod reliability was assessed by deattenuated Pearson correlations between the FFQ and 24-hour recalls. Attenuation factors and calibration corrected effect estimates for bone density were calculated to determine the potential influence of measurement error on associations with health outcomes.

Results The ICC (0.61) for repeat administrations and deattenuated Pearson correlation between the FFQ and 24-hour recalls (r=0.60) for all subjects indicated reproducibility and intermethod reliability (Pearson r=0.50 to 0.74 across sex and age groups). Attenuation factors were ≤0.50 for all sex and age groups and lower for non-Hispanic blacks (λ=0.20) and Hispanics (λ=0.26) than for non-Hispanic whites (λ=0.42).

Conclusions The Bone Mineral Density in Childhood Study calcium FFQ appears to provide a useful tool for assessing calcium intake in children and adolescents drawn from multiracial/ethnic populations and/or spanning a wide age range. However, similar to other FFQs, attenuation factors were substantially <1, indicating the potential for appreciable measurement error bias. Calibration correction should be performed and racial/ethnic differences in performance considered when analyzing and interpreting findings based on this instrument.
J Acad Nutr Diet. 2015;115:519-527.

FOOD FREQUENCY QUESTIONNAIRES (FFQS) HAVE been used extensively to obtain information on usual intake of foods and nutrients, including calcium, and to examine estimates of calcium intake in relation to health outcomes.[1-6] Advantages of FFQs for measuring dietary calcium intake, especially in large studies, include low cost, ease of administration, and focus on usual consumption of calcium-containing foods and supplements. Despite these strengths, nutrient intakes estimated from FFQs are subject to measurement error due to the limited number of food items queried, reliance on generic memory, potential reporting biases, and inaccuracies in food and nutrient databases. To properly interpret findings from a newly developed FFQ, reproducibility (intramethod reliability) and validation studies must be conducted. Ideally, the extent of error is determined from objective measures of true dietary intake; however, such measures are rare and confined to a select few recovery biomarkers.[7] Therefore, intermethod reliability studies are generally conducted that compare the performance of the new instrument with an accepted method. These studies are often referred to as relative validation studies, although an accurate measure of dietary intake is required to determine validity; therefore, the term *intermethod reliability* is preferable.[8] Multiple, unannounced 24-hour dietary recalls or dietary records are commonly used to assess intermethod reliability under the assumption that these methods provide an unbiased measure of dietary intake for a single day.[9-12] Regression calibration provides a means to correct for measurement error in the primary instrument (ie, the FFQ) and to obtain relatively unbiased measures of association between a food or nutrient and health outcome when the assumptions for proper reference instrument are met.[10,13,14]

The Bone Mineral Density in Childhood Study (BMDCS) was a large, national cohort of US children and adolescents designed to develop reference data for bone mass and density and to investigate the determinants of bone accretion.[15,16] To examine associations between calcium intake and bone mineral accrual in this cohort, a dietary assessment instrument designed for use in a nationally representative pediatric population was required. Therefore, an FFQ developed from dietary data collected as part of the third National Health and Nutrition Examination Study (NHANES III) and tailored for pediatric use was designed by NutritionQuest and used as the primary instrument to assess calcium intake. The purpose of this study was to examine the reproducibility and intermethod reliability of the BMDCS calcium FFQ for estimating usual dietary calcium intake in representative subsets of children and adolescents participating in the BMDCS. In addition, regression calibration was performed to assess potential measurement error bias when using the BMDCS calcium FFQ and to obtain correction factors that may be used to inform future studies. An example is provided investigating the impact of measurement error when estimating associations of calcium intake with bone mineral content (BMC) and areal bone mineral density (aBMD) at the spine to highlight the extent to which measurement error may influence effect estimates (ie, regression coefficients) and whether calibration correction may provide results more in line with calcium supplementation trials.[17-19]

MATERIALS AND METHODS

Study Population

The BMDCS comprised approximately 2,000 participants recruited from five clinical centers in the United States: Children's Hospital of Los Angeles (Los Angeles, CA), Cincinnati Children's Hospital Medical Center (Cincinnati, OH), Creighton University (Omaha, NE), Children's Hospital of Philadelphia (Philadelphia, PA), and Columbia University (New York, NY). The initial cohort consisted of a multiracial/ethnic sample of girls aged 6 to15 years and boys aged 6 to 16 years who were enrolled between July 2002 and November 2003 and followed annually for 6 years.[15] A second recruitment period, including individuals aged 5 and 19 years, occurred between August 2006 and November 2007 to increase the number of younger and older participants.[16] Healthy, normally developing children and adolescents were enrolled in the study. Participant recruitment and enrollment criteria have been described previously.[15,16] Informed consent was obtained from participants aged 18 years or older. Informed consent was obtained from the participant's parent or guardian and assent obtained from the participant for those younger than age 18 years. The study protocol was approved by the institutional review boards of all five centers.

Two substudies were conducted (reproducibility and intermethod reliability), each using a stratified, random sample of BMDCS participants aged 5 to 21 years of age at enrollment into the substudies. The goal of the reproducibility substudy was to assess the test–retest performance of the FFQ. Recruitment of participants (n=69) occurred from 2006 to 2007 by randomly selecting participants within strata defined by sex, age (≤13 years and >13 years), and center. The first FFQ (FFQ_1) was administered during the visit. The questionnaire was completed by study participants older than age 13 years and by the parent and child together for those aged ≤13 years. The second FFQ (FFQ_2) was mailed to participants 1 to 2 weeks after the study visit and returned within 1 month. Visual aids, including pictures with glasses and bowls and lines indicating volumes, were provided to aid participants in assessing food quantity.

The goal of the intermethod reliability substudy was to assess the accuracy of the FFQ in relation to three unannounced, interviewer-administered 24-hour dietary recalls. Recruitment of participants (n=430) occurred from 2006 to 2009. Participants were randomly selected within strata defined by sex, age (≤13 years and >13 years), and center. The sampling frame provided a subsample generally representative of the BMDCS cohort. The FFQ was administered during a clinical visit occurring from 2006 to 2009. The three unannounced 24-hour recalls were conducted via telephone by a study registered dietitian nutritionist (RDN) within 1 month of completing the FFQ. Visual aids, including pictures with glasses and bowls and lines indicating volumes, were provided to assist in portion size estimation. The FFQ and 24-hour recalls were completed by study participants older than age 13 years and by the parent and child together for those aged 13 years or younger. A total of 393 participants completed all three recalls and were included in our analyses. Compensation of $10 was provided for participation in a substudy.

Development of the FFQ

The BMDCS FFQ was designed to measure food items contributing calcium to the diets of children and adolescents during the previous 7 days. The FFQ was developed by NutritionQuest in accordance with procedures for the Block dietary questionnaires and screeners.[20-22] The initial list of calcium-containing foods was obtained from 24-hour recalls collected as part of NHANES III.[23] Individual food items were then aggregated into categories based on similarities in food type, nutrient content, and culinary use, and rank-ordered on contribution to total calcium intake. The minimum set of food categories that could capture ≥90% of calcium intake in each of nine prespecified age-race/ethnicity groups (ie, 6 to 10 years, 11 to 13 years, and 14 to 19 years; and Hispanic, non-Hispanic black, and non-Hispanic white) were included on the FFQ. The final questionnaire included queries on 45 distinct food and beverage categories (see the Figure, available online at www.andjrnl.org). Five responses were provided to assess how often the foods and beverages were consumed during the preceding week (1 day/wk, 2 days/wk, 3 to 4 days/wk, 5 to 6 days/wk, or every day). In addition, four portion size options were provided for each category to capture usual amounts consumed. Dietary calcium intake (in milligrams per day) was estimated from the questionnaire using food composition tables developed and maintained by NutritionQuest. For our analysis, instrument performance for calcium intake from foods alone was examined because quantitative information on supplement use for participants in the substudies was not collected on the FFQ. The calcium-specific FFQ used in the BMDCS is available from NutritionQuest (http://nutritionquest.com/).

Administration of the 24-Hour Recalls

Three nonconsecutive, unannounced 24-hour recalls were conducted via telephone by trained RDNs from Cincinnati Children's Hospital Medical Center and Children's Hospital of Philadelphia. Cincinnati Children's Hospital Medical Center was responsible for interviews with participants in Cincinnati, Creighton, and the male participants in Los Angeles. Children's Hospital of Philadelphia was responsible for interviews with participants in Philadelphia, New York, and the female participants in Los Angeles. The interviewer asked the participant about the foods and beverages consumed during the previous 24-hours employing a multiple-pass method to enhance detail and limit omissions. Quantities were expressed in terms of typical bowls, cups, and so on. Study participants were encouraged to consult the food-model pamphlet distributed at the study visit to assist in assessing portion sizes. The recalls were planned to occur on 2 weekdays and 1 weekend day. The recalls were collected and analyzed using the Nutrition Data System for Research software versions 2006-2009, developed by the Nutrition Coordinating Center at the University of Minnesota.

Bone Densitometry

Dual-energy x-ray absorptiometry scans were performed using a dedicated Hologic, Inc, bone densitometer (QDR4500A, QDR4500W, Apex, and Delphi A models) at each clinical center. Scans were performed for each participant according to the manufacturer's guidelines to obtain spine BMC and aBMD. Calibration stability across sites and over time was maintained using traveling and site-specific phantoms as previously described.[15] All scans were analyzed by the Dual-Energy X-Ray Absorptiometry Core Laboratory at the University of California, San Francisco, using Hologic software release 12.3 (2004) to obtain BMC and aBMD values for each participant. Data from the lumbar spine scans were used in our analysis as they have shown associations with calcium intake in the BMDCS cohort,[24] and the lumbar spine is a recommended site for pediatric bone health assessment.[25]

Additional Measures

Weight (0.1 kg) was measured on a digital scale and height (0.1 cm) was measured using a stadiometer. Sexual maturation was assessed by physical exam of testes volume (male participants) and breast size (female participants) by a pediatric endocrinologist or nurse practitioner. Sexual maturity was defined according to the Tanner criteria. Usual weight-bearing physical activity (in hours per week) was assessed with the modified self-report tool.[26] The questionnaire listed 37 weight-bearing activities, such as walking, basketball, dance, and so on, in which children and adolescents are likely to participate. Participants indicated the time spent in activities in which they participated during the previous week.

Statistical Analysis

Reproducibility. The mean difference and 95% CI for calcium intake estimated from FFQ_1 and FFQ_2 and the intraclass correlation (ICC) were used to assess reproducibility (intramethod reliability). High ICC values, denoting low within-subject variation, reflect a high degree of reproducibility. Results were calculated for all participants and for younger (aged 5 to 13 years) and older (aged 14 to 21 years) participants separately. ICC$\geq$0.6 was considered a priori to reflect appreciable reproducibility.

Intermethod Reliability. The mean difference and 95% CI for calcium intake estimated from the average of three 24-hour recalls and the FFQ, deattenuated Pearson correlations, and attenuation factors were used assess intermethod reliability. Deattenuation corrects for random measurement error (ie, day-to-day variability in the 24-hour recalls) in the observed correlation.[7] To calculate the deattenuated Pearson correlation (r_t), the following formula was used: $r_t = r_o \sqrt{(1+[S^2_w/S^2_b]/n_x)}$ where r_o is the observed correlation, S^2_w/S^2_b is ratio of the within- to between-subject variances in calcium intake estimated from repeat measures, and n_x is the number of 24-hour recalls. Deattenuated correlations in the range of r=0.4 to r=0.7 were considered a priori to reflect intermethod reliability.[9,27] Attenuation factors, calculated as the slope of the regression of the average of the three 24-hour recalls on the FFQ, provide a measure of the expected bias due to measurement error in the FFQ.[28] The attenuation factor (λ) is a multiplicative scalar that operates on the true value biasing it to the extent denoted by λ, where, for example, true value=observed value/λ. Attenuation factors closer to one indicate minimal bias, whereas lower values denote greater expected attenuation when estimating associations between dietary intake and health parameters. Calcium intakes were natural log transformed where necessary to better meet model assumptions. Results were obtained for all participants and for predefined sex-age and race/ethnic groups.

Calibration Correction. Uncorrected (referred to herein as naïve) and calibration corrected multiple linear regression was used to examine associations between usual calcium intake and BMC and aBMD at the spine adjusted for age (continuous), sex, race (black vs nonblack), height (continuous), total physical activity (continuous), tanner stage, and clinical center. Calibration was performed using the method described by Spiegelman and colleagues.[29] for linear regression. Calcium intake assessed via the FFQ was assumed to be measured with error and the average of the three 24-hour recalls was used as the reference instrument. All covariates were assumed to be measured without error and all regressions were performed on the log-log scale to better meet model assumptions.

RESULTS

The mean age of participants in the BMDCS reproducibility and intermethod reliability substudies was 14.5 years (minimum=5 years, maximum=21 years) and 13.1 years (minimum=5 years, maximum=21 years), respectively, at the time of data collection (Table 1). For the reproducibility subsample, non-Hispanic whites comprised the largest proportion of participants (52%), followed by Hispanics, non-Hispanic blacks, and Asian Americans with the majority of participants at Tanner stage 4 or 5. For the intermethod reliability subsample, non-Hispanic whites comprised the largest proportion of participants (49%), followed by non-Hispanic blacks, Hispanics, and Asian Americans. The greatest proportions of participants were Tanner stage 1 and 5 with fewer at intermediate stages. Approximately 45% of participants in both samples were male with mean heights of 159 cm and 151 cm in the reproducibility and intermethod reliability subsamples, respectively.

Reproducibility

The mean difference in calcium intake estimated with repeat administrations of the FFQ (FFQ_1-FFQ_2) was 17 mg/day (95% CI −89 to 123) (Table 2). The ICC for calcium intake (log scale) indicated reproducibility of the FFQ (ICC=0.61). In analyses stratified on age, there was no difference in mean calcium intake upon repeat administration in both age groups; however, the ICC was lower for younger (aged 5 to 13 years) when compared with older (aged 14 to 21 years) participants. The low correlation for younger children was highly influenced by one outlying value (Cook's D=2.2) and improved to ICC=0.52 upon removal of that single value.

Table 1. Characteristics of participants in the Bone Mineral Density in Childhood Study calcium food frequency questionnaire reproducibility and intermethod reliability substudies[a]

Characteristic	Reproducibility Substudy (n=69)	Intermethod Reliability Substudy (n=393)
	←mean±standard deviation→	
Age (y)	14.5±3.9	13.1±5.2
Height (cm)	158.5±17.8	150.5±22.3
	←n (%)→	
Male participants	31 (44.9)	184 (46.8)
Race/ethnicity		
Asian	8 (11.6)	20 (5.1)
Non-Hispanic black	10 (14.5)	81 (20.6)
Hispanic	12 (17.4)	68 (17.3)
Non-Hispanic white	36 (52.2)	194 (49.4)
Other	3 (4.4)	30 (7.6)
Tanner stage		
1	8 (11.6)	125 (31.8)
2	4 (5.8)	12 (3.1)
3	10 (14.5)	18 (4.6)
4	12 (17.4)	38 (9.7)
5	33 (47.8)	170 (43.3)
Missing	2 (2.9)	30 (7.6)

[a]Values may not sum to total due to missing data.

Intermethod Reliability

The FFQ underestimated calcium intake relative to the average of three 24-hour recalls (Table 3). The mean difference for $calcium_{recalls}-calcium_{FFQ}$ was 72 mg/day (95% CI 12 to 131) for all subjects indicating a systematic bias at the group level as expected based on the FFQ development protocol. CIs for the mean difference between the FFQ and 24-hour recalls were wide for all age-sex and race subgroups, but reached statistical significance for all females, females aged 10 to 14 years, females aged 15 to 21 years, and non-Hispanic whites. The deattenuated Pearson correlation coefficient for the FFQ and average of three 24-hour recalls was r=0.60 for all subjects indicating intermethod reliability and ranged from r=0.50 to r=0.74 across sex and age groups. Correlations were modestly higher for females than for males and for participants aged 5 to 7 years. In race/ethnic-specific analyses, the correlations were lowest for non-Hispanic blacks (r=0.48), followed by Hispanics (r=0.54) and non-Hispanic whites (r=0.62). The attenuation factor for the FFQ relative to the average of three 24-hour recalls was λ=0.36 for all subjects, indicating the extent to which associations between calcium intake estimated from the FFQ and health outcomes may be biased toward the null in univariate analyses. Attenuation factors were ≤0.50 for all sex and age groups and λ=0.20 for non-Hispanic blacks, λ=0.26 for Hispanics, and λ=0.42 for non-Hispanic whites.

Calibration Correction

Uncorrected and calibration corrected estimates for the association of calcium intake with BMC and aBMD at the spine are provided in Table 4. There were no statistically significant associations overall or for non-Hispanic black or Hispanic participants; however, the uncorrected results were biased toward no association to the degree expected by the attenuation factors. For non-Hispanic whites, in uncorrected models, a 10% increase in calcium intake was associated with a 0.54% increase in BMC (P=0.02) and a 0.41% increase in aBMD (P=0.03). Adjusted for measurement error, a 10% increase in calcium intake was associated with a 1.52% increase in BMC (P=0.03) and a 1.16% increase in aBMD (P=0.04). In our sample, an increase in calcium consumption of roughly 300 mg/day, equivalent to 1 cup milk or the low range of supplemental calcium provided in previous trials, corresponded to an approximate increase of 40% in daily intake. In calibration corrected models this translates into an approximate increase of 2.13% and 1.62% in BMC and aBMD, respectively, when estimated within the range of the data among non-Hispanic whites. Overall, attenuation factors obtained from the multivariable models were generally similar to those obtained for the univariate model presented in Table 3 (see also Table 5 [available online at www.andjrnl.org], where data for all subjects only is shown).

Table 2. Reproducibility of calcium intake for repeat administrations of the Bone Mineral Density in Childhood Study calcium food frequency questionnaire (FFQ_1 and FFQ_2)

Calcium intake (mg/d)	FFQ_1	FFQ_2[a]	Difference FFQ_1−FFQ_2	Intraclass correlation coefficient[b]
	←median (25th, 75th)→		←mean (95% CI)→	
All subjects (n=69)	655 (485, 883)	685 (417, 979)	17 (−89 to 123)	0.61
Ages 5-13 y (n=31)	744 (474, 909)	826 (557, 984)	−87 (−235 to 61)	0.31
Ages 14-21 y (n=38)	636 (484, 739)	539 (398, 907)	102 (−48 to 252)	0.75

[a]FFQ_2 was completed and returned ∼1 month after FFQ_1.
[b]Calculated for the natural log transformed values.

DISCUSSION

These data indicate that the BMDCS calcium FFQ provides reasonably reproducible and reliable estimates of calcium intake in representative samples of children and adolescents participating in the BMDCS when compared with the average of three 24-hour dietary recalls. Calcium intake was generally underestimated by the FFQ relative to the 24-hour recalls; however, correlational analyses demonstrated deattenuated associations consistent with our a priori criterion for intermethod reliability. Despite reasonable agreement between instruments, attenuation factors were ≤0.50 across all subgroups, highlighting the potential for substantial measurement error bias when estimating associations between calcium intake and health outcomes. Nutrition professionals using the BMDCS calcium FFQ are, therefore, recommended to obtain calibration corrected estimates, in addition to naïve estimates that assume no measurement error, when assessing associations between calcium intake and health outcomes to better understand the extent to which measurement error may have influenced the findings.

The BMDCS FFQ was able to reproducibly measure dietary calcium intake. Previous research assessing the reproducibility of both paper and web-based FFQs designed specifically to measure dietary calcium intake in children and adolescents have reported correlations for test−retest performance ranging from r=0.68 to r=0.79.[6,30-32] The modestly weaker correlation in our sample was likely due to differences in the time frame queried by the FFQ. Unlike previous FFQs designed to assess calcium intake during the preceding

Table 3. Intermethod reliability of the Bone Mineral Density in Childhood Study calcium food frequently questionnaire (FFQ) to measure calcium intake when compared with the average of three 24-hour diet recalls[a]

Calcium intake (mg/d)	Diet recalls	FFQ	Difference recalls−FFQ	Deattenuated Pearson r[bd]	Attenuation factor (λ)[cd]
	←median (25th, 75th)→		←mean (95% CI)→		
All subjects (n=393)	871 (652, 1,160)	773 (520, 1,145)	72 (12 to 131)	0.60	0.36
Males (n=184)	924 (741, 1,327)	891 (624, 1,269)	39 (−70 to 149)	0.55	0.35
Ages 5-7 y (n=56)	940 (756, 1,112)	997 (689, 1,188)	0 (−100 to 100)	0.74	0.37
Ages 10-14 y (n=62)	1000 (812, 1,330)	917 (669, 1,262)	2 (−133 to 138)	0.54	0.29
Ages 15-21 y (n=66)	845 (656, 1,424)	776 (553, 1,352)	108 (−161 to 376)	0.50	0.37
Females (n=209)	831 (605, 1,116)	650 (458, 993)	100 (42 to 157)	0.61	0.34
Ages 5-7 y (n=57)	916 (637, 1,157)	821 (542, 1,127)	59 (−26 to 145)	0.69	0.50
Ages 10-14 y (n=73)	874 (671, 1,111)	639 (491, 964)	126 (20 to 232)	0.54	0.27
Ages 15-21 y (n=79)	743 (538, 1,020)	593 (385, 905)	105 (3 to 207)	0.61	0.30
Ethnicity					
Black (n=81)	778 (633, 900)	706 (478, 925)	28 (−84 to 140)	0.48	0.20
Hispanic (n=68)	835 (602, 1,132)	700 (467, 1,075)	42 (−99 to 183)	0.54	0.26
White (n=194)	963 (744, 1,280)	872 (557, 1,172)	113 (21 to 205)	0.62	0.42

[a]All values for the 24-hour diet recall are for the 3-day average.
[b]Deattenuated for random measurement error in the 24-hour diet recalls: $r_o\sqrt{(1+[S^2_w/S^2_b]/3)}$, where r_o=observed Pearson correlation.
[c]λ=Slope of the regression of the average of the three 24-hour diet recalls on the FFQ.
[d]Calculated for the natural log transformed values.

Table 4. Associations of calcium intake with bone mineral content and areal bone mineral density at the spine for participants in the Bone Mineral Density in Childhood Study intermethod reliability substudy[a]

	Bone Mineral Content		Bone Mineral Density	
Characteristic	**β±standard error**	***P* value**	**β±standard error**	***P* value**
All subjects (n=358)				
FFQ[b]				
Uncorrected	.016±.014	0.27	.011±.012	0.36
Calibration corrected[c]	.051±.047	0.28	.035±.038	0.36
Black (n=76)				
FFQ				
Uncorrected	.018±.030	0.55	.001±.025	0.97
Calibration corrected[c]	.110±.192	0.57	.006±.153	0.97
Hispanic (n=64)				
FFQ				
Uncorrected	−.055±.028	0.05	−.032±.022	0.15
Calibration corrected[c]	−.228±.144	0.11	−.134±.105	0.20
White (n = 175)				
FFQ				
Uncorrected	.054±.023	0.02	.041±.019	0.03
Calibration corrected[c]	.152±.069	0.03	.116±.057	0.04

[a]Associations for natural log-log models. All models adjusted for age (continuous), sex, race (black vs non-black), height (continuous), total physical activity (continuous), Tanner stage, and clinical center in the calibration and health parameter models where appropriate. Subjects who refused Tanner staging (n=30) or did not undergo bone mineral content/density testing (n=5) were excluded from analyses.
[b]FFQ=food frequency questionnaire.
[c]Calibration correction performed using the method described by Spiegelman and colleagues[29] for linear regression.

month, the BMDCS calcium FFQ queries intake occurring during the preceding week. This decision was made to limit cognitive difficulties in estimating the usual frequency and quantity of foods items consumed over longer time spans in an attempt to improve accuracy. Thus, the repeat administrations in our study do not reflect dietary intake occurring during the same time period with an expectant reduction in the ICC due to incorporating true within-person variation in calcium intake. The reason for the discrepancy in the ICC for younger (aged 5 to 13 years) when compared with older (aged 14 to 20 years) participants is unknown, but could reflect a more varied diet of calcium-containing foods in younger participants or simply sampling variation. Differences in test–retest performance according to age in previous reports have been inconsistent.[6,30,31]

Compared with the average of three unannounced 24-hour recalls, the BMDCS FFQ underestimated dietary calcium intake on average. This was expected because the FFQ was designed to capture ≥90% of calcium intake in select age-race/ethnicity groups based on 24-hour recall data. Previous studies have reported lower point estimates for mean calcium intake estimated from a short 10-item FFQ relative to the average from 24-hour recalls[31] and higher mean calcium intake estimated from a 41-item FFQ relative to a 7-day weighted food record.[6] Numerous factors influence the assessment of absolute calcium intake on an FFQ, including the length of the questionnaire, food list, response options, nutrient calculations, and subject instructions[7] and must be considered along with the study objectives and participant burden when designing the questionnaire. The deattenuated correlation for the BMDCS FFQ and 24-hour recalls of r=0.60 was similar to, or better than, correlations with reference measures for other FFQs designed to measure calcium intake in children and adolescents in the United States[6,30-33] and other countries[34-38] despite the wide age range and multiracial/ethnic sample. This highlights the potential utility of this instrument for assessing calcium intake in future studies drawn from multiracial/ethnic populations or those spanning a wide age range. The higher correlation for female participants than for male participants in our sample has not been consistently observed in previous reports, and the lack of uniformity across age groups is in agreement with earlier studies.[6,30,31] A noteworthy finding was that despite developing the FFQ from items providing the greatest contribution to dietary calcium intake in Hispanic, non-Hispanic black, non-Hispanic white children and adolescents, correlations between the FFQ and 24-hour recalls were lower for non-Hispanic blacks and Hispanics than for non-Hispanic whites. Similar differences have been reported previously.[6] The reason for this discrepancy is unknown and highlights unresolved challenges in culturally tailoring aspects of the questionnaire beyond the list of foods to elicit more accurate responses in specific racial/ethnic groups.

Attenuation factors obtained from the slope of the regression of the 24-hour recalls on the FFQ indicated the potential for appreciable measurement error bias when estimating associations between calcium intake and health outcomes. Examples for the association of dietary calcium intake with BMC and aBMD at the spine demonstrated naïve associations were attenuated by approximately two-thirds relative to calibration-corrected results. Measurement error of this magnitude has the potential to severely distort associations between calcium intake and health outcomes with the direction of the bias uncertain in multivariable statistical models.[28,39,40] Furthermore, attenuation factors were lower for Hispanics and non-Hispanic blacks than for non-Hispanic whites, indicating that naïve effect estimates may be more biased for these groups. Measurement error also influences study power, reducing effective sample size by a factor inversely proportional to λ^2.[28] Thus, the ability to detect statistically significant associations in Hispanics and non-Hispanics blacks, relative to non-Hispanic whites, would also be reduced. Calcium supplementation trials in children and adolescents have generally reported modest increases (1% to 6%) in BMC and aBMD during active supplementation, although it is unclear whether improvements are confined to children with low baseline levels or persist beyond the end of supplementation.[17-19] In calibration-corrected models, a $\sim$300 mg/day increase in calcium intake was related to a 2.11% and 1.64% increase in BMC and aBMD in non-Hispanic whites, more in line with results obtained from calcium supplementation trials. Calibration correction using multiple 24-hour recalls as the reference instrument may allow for more accurate results because effect estimates have been, in part, corrected for measurement error. A measure of true usual calcium intake would be required to eliminate bias due to measurement error. Although traditional regression calibration provides a means to correct naïve effect estimates (eg, regression coefficients), it does not recapture lost study power, improve tests for statistical significance, and is influenced by error in the assessment of model covariates.

Regression calibration also relies on the assumption that the error in the reference instrument is unbiased and contains only within-person error that is uncorrelated with errors in the primary instrument.[9] Comparisons of self-report instruments with recovery biomarkers have demonstrated both intake-related and correlated person-specific biases violating both assumptions of a proper reference instrument.[13,14] To the extent that measurement error in the 24-hour dietary recall method is associated with error in the BMDCS FFQ, correlational analyses (ie, intermethod reliability) and attenuation factors will have been overestimated and the true attenuation for estimated effects greater than that reported. Despite these limitations, regression calibration using the 24-hour recall as the reference instrument is preferable to performing no adjustment because improvements in parameter estimates have been shown for both univariate and multivariable methods.[41] Nutrition and dietetics practitioners using the BMDCS calcium FFQ are therefore advised to report the naïve estimates that assume no measurement error as well as calibration-corrected results when possible.

Strengths of this study include the rigorous methodology employed in the development of the FFQ, the large sample size for the intermethod reliability substudy, the RDN-administered 24-hour dietary recalls, and the multiracial/ethnic composition of the substudies contributing to the generalizability of the results. There were also limitations. First, as previously discussed, correlations and attenuation factors for the BMDCS FFQ presented here may be overly optimistic relative to true usual calcium intake due to correlated errors in the primary and reference instruments. However, this limitation is common to all dietary intermethod reliability studies in which an unbiased measure of dietary intake is unavailable. Second, the study power to detect small and small-to-moderate mean differences between instruments was limited for the intermethod reliability and reproducibility substudies, respectively. Third, timing with respect to the administration of the FFQ and 24-hour recalls may have modestly reduced correlations between measures because they did not assess dietary intake occurring over the exact same period of time. Fourth, we were unable to examine whether calcium intake from supplements had any material influence on measures of reproducibility and intermethod reliability. Fifth, Tanner staging was not calibrated across clinicians; however, Tanner stage had little influence on the regression calibration or health outcome models, and any misclassification is, therefore, not expected to have materially influenced the study findings. In addition, the BMDCS calcium FFQ was developed using 24-hour recalls collected during NHANES III, warranting examination of the generalizability of the tool for capturing contemporary food and beverage intakes. The food items included on the FFQ remained capable of capturing >90% of calcium intake among US children aged 8 to 17 years based 24-hour recall data collected during the 2005-2006 NHANES (personal communication, NutritionQuest, October 2014). Thus, similar performance of the instrument is expected for contemporary nutrition research studies.

CONCLUSIONS

The BMDCS calcium FFQ appears to provide a useful tool for assessing dietary calcium intake in children and adolescents. The ability to administer a single, short instrument with demonstrated performance across a wide age range and in racially/ethnically heterogeneous populations makes the instrument an appealing option for nutrition and dietetics practitioners. While meeting our a priori criterion for intermethod reliability in all racial/ethnic groups examined, the performance of the questionnaire was found to vary. Therefore, consideration with respect to potential racial/ethnic differences in measurement error bias needs to be given when analyzing and interpreting findings from naïve analyses based on this instrument. When possible, internal calibration studies should be conducted to facilitate the calculation of calibration corrected effect estimates. In the absence of an internal calibration study, nutrition and dietetics practitioners may also consider using our findings to conduct external calibration sensitivity analyses to account for measurement error.

References

1. Feskanich D, Willett WC, Colditz GA. Calcium, vitamin D, milk consumption, and hip fractures: A prospective study among postmenopausal women. *Am J Clin Nutr*. 2003;77(2):504-511.
2. Park SY, Murphy SP, Wilkens LR, Stram DO, Henderson BE, Kolonel LN. Calcium, vitamin D, and dairy product intake and

prostate cancer risk: The Multiethnic Cohort Study. *Am J Epidemiol*. 2007;166(11):1259-1269.

3. Xiao Q, Murphy RA, Houston DK, Harris TB, Chow WH, Park Y. Dietary and supplemental calcium intake and cardiovascular disease mortality: The National Institutes of Health-AARP diet and health study. *JAMA Intern Med*. 2013;173(8):639-646.
4. Black RE, Williams SM, Jones IE, Goulding A. Children who avoid drinking cow milk have low dietary calcium intakes and poor bone health. *Am J Clin Nutr*. 2002;76(3):675-680.
5. Rockett HR, Wolf AM, Colditz GA. Development and reproducibility of a food frequency questionnaire to assess diets of older children and adolescents. *J Am Diet Assoc*. 1995;95(3):336-340.
6. Zemel BS, Carey LB, Paulhamus DR, Stallings VA, Ittenbach RF. Quantifying calcium intake in school age children: Development and validation of the Calcium Counts! food frequency questionnaire. *Am J Hum Biol*. 2010;22(2):180-186.
7. Willett W. *Nutritional Epidemiology*. 3rd ed. New York, NY: Oxford University Press; 2013.
8. White E, Armstrong BK, Saracci R, Armstrong BK. *Principles of Exposure Measurement in Epidemiology: Collecting, Evaluating, and Improving Measures of Disease Risk Factors*. 2nd ed. New York, NY: Oxford University Press; 2008.
9. Subar AF, Thompson FE, Kipnis V, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: The Eating at America's Table Study. *Am J Epidemiol*. 2001;154(12):1089-1099.
10. Stram DO, Hankin JH, Wilkens LR, et al. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. *Am J Epidemiol*. 2000;151(4):358-370.
11. Horn-Ross PL, Lee VS, Collins CN, et al. Dietary assessment in the California Teachers Study: Reproducibility and validity. *Cancer Causes Control*. 2008;19(6):595-603.
12. Carithers TC, Talegawkar SA, Rowser ML, et al. Validity and calibration of food frequency questionnaires used with African-American adults in the Jackson Heart Study. *J Am Diet Assoc*. 2009;109(7): 1184-1193.
13. Kipnis V, Carroll RJ, Freedman LS, Li L. Implications of a new dietary measurement error model for estimation of relative risk: Application to four calibration studies. *Am J Epidemiol*. 1999;150(6): 642-651.
14. Kipnis V, Midthune D, Freedman LS, et al. Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am J Epidemiol*. 2001;153(4):394-403.
15. Kalkwarf HJ, Zemel BS, Gilsanz V, et al. The bone mineral density in childhood study: Bone mineral content and density according to age, sex, and race. *J Clin Endocrinol Metab*. 2007;92(6):2087-2099.
16. Zemel BS, Kalkwarf HJ, Gilsanz V, et al. Revised reference curves for bone mineral content and areal bone mineral density according to age and sex for black and non-black children: Results of the bone mineral density in childhood study. *J Clin Endocrinol Metab*. 2011;96(10):3160-3169.
17. Lanou AJ, Berkow SE, Barnard ND. Calcium, dairy products, and bone health in children and young adults: A reevaluation of the evidence. *Pediatrics*. 2005;115(3):736-743.
18. Winzenberg T, Shaw K, Fryer J, Jones G. Effects of calcium supplementation on bone density in healthy children: Meta-analysis of randomised controlled trials. *BMJ*. 2006;333(7572):775.
19. Huncharek M, Muscat J, Kupelnick B. Impact of dairy products and dietary calcium on bone-mineral content in children: Results of a meta-analysis. *Bone*. 2008;43(2):312-321.
20. Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, Gardner L. A data-based approach to diet questionnaire design and testing. *Am J Epidemiol*. 1986;124(3):453-469.
21. Block G, Coyle LM, Hartman AM, Scoppa SM. Revision of dietary analysis software for the Health Habits and History Questionnaire. *Am J Epidemiol*. 1994;139(12):1190-1196.
22. Block G, Gillespie C, Rosenbaum EH, Jenson C. A rapid food screener to assess fat and fruit and vegetable intake. *Am J Prev Med*. 2000;18(4):284-288.
23. Centers for Disease Control and Prevention, National Center for Health Statistics. National Health and Nutrition Examination Survey data. 1994. http://www.cdc.gov/nchs/nhanes/nh3data.htm. Accessed June 24, 2014.
24. Lappe JM, Watson P, Gilsanz V, et al. The longitudinal effects of physical activity and dietary calcium on bone mass accrual across stages of pubertal development. *J Bone Miner Res*. 2015;30(1):156-164.
25. Gordon CM, Leonard MB, Zemel BS. International Society for Clinical D. 2013 Pediatric Position Development Conference: Executive summary and reflections. *J Clin Densitom*. 2014;17(2):219-224.
26. Slemenda CW, Miller JZ, Hui SL, Reister TK, Johnston CC Jr. Role of physical activity in the development of skeletal mass in children. *J Bone Miner Res*. 1991;6(11):1227-1233.
27. Thompson FE, Byers T. Dietary assessment resource manual. *J Nutr*. 1994;124(11 suppl):2245S-2317S.
28. Schatzkin A, Kipnis V, Carroll RJ, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: Results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. *Int J Epidemiol*. 2003;32(6):1054-1062.
29. Spiegelman D, McDermott A, Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr*. 1997;65(4 suppl):1179S-1186S.
30. Jensen JK, Gustafson D, Boushey CJ, et al. Development of a food frequency questionnaire to estimate calcium intake of Asian, Hispanic, and white youth. *J Am Diet Assoc*. 2004;104(5):762-769.
31. Harnack LJ, Lytle LA, Story M, et al. Reliability and validity of a brief questionnaire to assess calcium intake of middle-school-aged children. *J Am Diet Assoc*. 2006;106(11):1790-1795.
32. Wong SS, Boushey CJ, Novotny R, Gustafson DR. Evaluation of a computerized food frequency questionnaire to estimate calcium intake of Asian, Hispanic, and non-Hispanic white youth. *J Am Diet Assoc*. 2008;108(3):539-543.
33. Rockett HR, Breitenbach M, Frazier AL, et al. Validation of a youth/adolescent food frequency questionnaire. *Prev Med*. 1997;26(6): 808-816.
34. Taylor RW, Goulding A. Validation of a short food frequency questionnaire to assess calcium intake in children aged 3 to 6 years. *Eur J Clin Nutr*. 1998;52(6):464-465.
35. Bertoli S, Petroni ML, Pagliato E, et al. Validation of food frequency questionnaire for assessing dietary macronutrients and calcium intake in Italian children and adolescents. *J Pediatr Gastroenterol Nutr*. 2005;40(5):555-560.
36. Huybrechts I, De Bacquer D, Matthys C, De Backer G, De Henauw S. Validity and reproducibility of a semi-quantitative food-frequency questionnaire for estimating calcium intake in Belgian preschool children. *Br J Nutr*. 2006;95(4):802-816.
37. Magkos F, Manios Y, Babaroutsi E, Sidossis LS. Differences in the quantitative and qualitative performance of a calcium-specific food frequency questionnaire across age and sex. *J Hum Nutr Diet*. 2006;19(5):331-342.
38. Moore M, Braid S, Falk B, Klentrou P. Daily calcium intake in male children and adolescents obtained from the rapid assessment method and the 24-hour recall method. *Nutr J*. 2007;6:24.
39. Fuller WA. *Measurement Error Models*. Hoboken, NJ: Wiley-Interscience; 2006.
40. Carroll RJ, Carroll RJ. *Measurement Error in Nonlinear Models: A Modern Perspective*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC; 2006.
41. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. *J Natl Cancer Inst*. 2011;103(14):1086-1092.

AUTHOR INFORMATION

N. J. Ollberding is an assistant professor, Division of Biostatistics and Epidemiology, and H. J. Kalkwarf is a professor, Division of General and Community Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH. V. Gilsanz is a professor, Division of Radiology, Children's Hospital of Los Angeles, Los Angeles, CA. J. M. Lappe is a professor, Division of Endocrinology, Creighton University, Omaha, NE. S. E. Oberfield is a professor, Division of Pediatric Endocrinology Diabetes and Metabolism, Columbia University, New York, NY. J. A. Shepherd is an associate adjunct professor, Department of Radiology, University of California, San Francisco. K. K. Winer is director, Pediatric Endocrinology Program, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD. B. S. Zemel is a professor, Division of Gastroenterology, Hepatology, and Nutrition, Children's Hospital of Philadelphia, Philadelphia, PA.

Address correspondence to: Nicholas J. Ollberding, PhD, Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Location S, MLC 5041, Cincinnati, OH 45229. E-mail: nicholas.ollberding@cchmc.org

STATEMENT OF POTENTIAL CONFLICT OF INTEREST

No potential conflict of interest was reported by the authors.

FUNDING/SUPPORT

This research was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (contracts N01-HD-1-3228, -3329, -3330, -3331, -3332, and -3333), the National Center for Research Resources (contracts UL1 RR026314 and UL1 RR024134), and General Clinical Research Grants (nos. M01-RR-000240 and M01-RR-08084).

ACKNOWLEDGEMENTS

The authors thank the research dietitians at Cincinnati Children's Hospital Medical Center and Children's Hospital of Philadelphia for their expert service and dedication to this project. The authors also thank the study members and their families for their participation.

Grains
Cold cereal
Hot cereal
Pancakes, waffles, or French toast
Granola bars or breakfast bars
Bread
Cornbread, corn muffin, or corn tortillas
Flour tortillas
Biscuits or muffins
Hamburger buns, hot dog buns, or bagels
Stuffing
Rice

Dairy
Milk
Milk on cereal
Cheese
Nachos with cheese
Yogurt
Sour cream
Ice cream or frozen yogurt

Combination foods
Hamburgers or cheeseburgers
Hot dogs or corn dogs
Tacos or burritos
Mixed dish with chicken (chicken and gravy or chicken pot pie)
Fish, fish sandwiches, or fish sticks
Mixed dish with beef (beef and noodles, hamburger mixed dishes, or stew)
Spaghetti, ravioli, or lasagna
Macaroni and cheese
Pizza or pizza pockets

Vegetables
Baked beans, refried beans, pinto beans, or red Mexican beans
Greens including spinach, mustard greens, or collards
Vegetable soup, tomato soup, or cream soup
Potatoes, including french fries, tater tots, or mashed or baked potatoes

Beverages
Sweetened carbonated beverages, any type
Fruit flavored drinks
Orange juice (queried calcium fortification)
Other fruit juices (queried calcium fortification)
Instant breakfast milkshakes
Milkshakes

Other
Eggs
Sausage
Gravy
Potato chips or other salty snacks
Cookies
Cake or cupcakes
Pudding
Chocolate candy

Figure. Food items included in the Bone Mineral Density in Childhood Study calcium food frequency questionnaire.

Table 5. Regression calibration model for natural log transformed calcium intake

All subjects (N=363)[a]	Intercept	FFQ[b] calcium[c]	Age	Male	Black	Hispanic	Other	Height	Physical activity
	←——— *mean±standard error*[d] ———→								
	6.894±0.102	0.307±0.037	0.001±0.01	0.057±0.045	−0.137±0.056	−0.057±0.071	−0.088±0.072	0.004±0.003	0.002±0.003
	Tanner 2	Tanner 3	Tanner 4	Tanner 5	Center A	Center B	Center D	Center E	
	−0.004±0.129	−0.081±0.127	−0.017±0.121	−0.202±0.134	−0.019±0.077	−0.132±0.079	−0.003±0.066	−0.02±0.073	

[a]The intercept represents the value for a participant at the mean of calcium intake (6.6 mg/d), age (13.1 y), and height (150.1 cm), who is female, non-Hispanic white, Tanner stage 1, reported engaging in no physical activity, and enrolled at clinical site C.

[b]FFQ=food frequency questionnaire.

[c]Attenuation factor (calibration coefficient).

[d]Values are the conditional mean±standard error obtained by regressing the average of the three 24-hour recalls on the FFQ+covariates.