Participant characteristics associated with errors in self-reported energy intake from the Women’s Health Initiative food-frequency questionnaire1–3

Neillann K Horner, Ruth E Patterson, Marian L Neuhouser, Johanna W Lampe, Shirley A Beresford, and Ross L Prentice

ABSTRACT

Background: Errors in self-reported dietary intake threaten inferences from studies relying on instruments such as food-frequency questionnaires (FFQs), food records, and food recalls.

Objective: The objective was to quantify the magnitude, direction, and predictors of errors associated with energy intakes estimated from the Women’s Health Initiative FFQ.

Design: Postmenopausal women (n = 102) provided data on sociodemographic and psychosocial characteristics that relate to errors in self-reported energy intake. Energy intake was objectively estimated as total energy expenditure, physical activity expenditure, and the thermic effect of food (10% addition to other components of total energy expenditure).

Results: Participants underreported energy intake on the FFQ by 20.8%; this error trended upward with younger age (P = 0.07) and social desirability (P = 0.09) but was not associated with body mass index (P = 0.95). The correlation coefficient between reported energy intake and total energy expenditure was 0.24; correlations were higher among women with less education, higher body mass index, and greater fat-free mass, social desirability, and dissatisfaction with perceived body size (all P < 0.10).

Conclusions: Energy intake is generally underreported, and both the magnitude of the error and the association of the self-reporting with objectively estimated intake appear to vary by participant characteristics. Studies relying on self-reported intake should include objective measures of energy expenditure in a subset of participants to identify person-specific bias within the study population for the dietary self-reporting tool; these data should be used to calibrate the self-reported data as an integral aspect of diet and disease association studies.


KEY WORDS Dietary records, systematic bias, dietary measurement error, energy expenditure, postmenopausal women, Women’s Health Initiative

INTRODUCTION

Nutrient intake estimates from self-reporting tools such as food-frequency questionnaires (FFQs), records, and recalls are known to contain considerable random error. Of greater concern is the accumulating evidence that these nutrient estimates also incorporate systematic error. In particular, dietary self-reporting appears to substantially underestimate energy intake, and there is some evidence that this underreporting varies by participant.
intake is critical for the interpretation of diet-disease relations in analyses that rely on self-reported dietary intake, and the development of a comprehensive measurement-error model has the potential to increase the validity and reliability of diet-disease association analysis.

SUBJECTS AND METHODS

Study overview

We recruited 102 postmenopausal women aged 50–79 y from the Seattle area through direct mailings by using the Washington State Department of Licensure list, flyers, and newspaper advertisements. Interested women contacted us by telephone, and we provided an overview of the study. Potential participants were screened for conditions that might interfere with body-composition calculations or biomarker measurements: bowel disease, diabetes or hypoglycemia, kidney disease, chronic lung disease, liver disease, claustrophobia, incontinence, systemic steroid medication use, weight change > 4.5 kg in the 2 mo before enrollment, and alcohol intake > 2 servings/d (24 g). None of the women enrolled in this study were participating in the WHI study.

Participants completed questionnaires and collected their urine for 24 h before each of 2 clinic visits scheduled 1 wk apart. At the visits, we received the 24-h urine collections, reviewed completed questionnaires, took body measurements, conducted indirect calorimetry, and drew blood. To measure activity, participants wore an accelerometer for 3 consecutive days during the week between clinic visits. The Institutional Review Board at the Fred Hutchinson Cancer Research Center approved all procedures.

Energy intake reporting on the FFQ

Before the clinic visits, participants received the WHI FFQ in the mail (visit 1) and in person (visit 2). Both FFQs were completed at home by the participants without assistance. A registered dietitian checked the FFQs for completeness at each visit. The WHI FFQ asks about usual dietary intake during the previous 3 mo and contains 3 sections: 1) 19 adjustment questions that are used in the analysis software to calculate the nutrient content of specific food items, 2) a listing of 122 foods or food groups with questions about the usual frequency of intake and portion size for each entry, and 3) 4 summary questions about the usual intake of fruit, vegetables, and fat added to foods and used in cooking. We analyzed the FFQ dietary data by using the University of Minnesota Nutrition Coordinating Center Nutrient Database (22), and the algorithms for analysis are described elsewhere (23). Compared with food records and recalls in a study of 113 WHI study participants, the completed WHI FFQs underestimated energy intakes by 8% and correlated with energy estimates from records or recalls at 0.4 (24). However, the WHI FFQ has not been validated against the doubly labeled water method.

Total energy expenditure

We objectively estimated energy intake by summing the components of TEE (RMR and AREE) with a 10% addition to this summation for the thermic effect of food (25). Below we provide details of our protocol and measures.

Resting metabolic rate

We measured RMR with a VMAX 2900 indirect calorimeter (SensorMedics, Loma Linda, CA). After 30 min for machine warm-up, volume calibration was conducted daily to ≤ 3% of 3 L according to the manufacturer’s recommendations. Gas calibrations were done before each subject was measured, with the use of 2 mixtures: 26% O2 and 0% CO2 and 20% O2 and 0.75% CO2.

We instructed participants to abstain from food and beverages, except water, for ≥ 8 h and to avoid strenuous activity for 48 h before each indirect calorimetry measurement. To begin the visit, participants rested quietly on a recliner for 30 min in the thermally neutral testing room. We explained the procedure and oriented participants to the equipment. The mixing chamber pump was turned on, and the plastic canopy was placed over the head and neck of the recumbent participant, with the vinyl skirt covering the torso. For the participants’ acclimation to the apparatus and our adjustment of pump speed, we allowed 2 min of data to expire before initiating formal data collection. We collected data points every 30 s and defined steady state as 10 min during which the volume of oxygen consumed, the minute ventilation, and the respiratory quotient did not vary by > 10%. If 10 min of steady state was achieved by 30 min of data collection, the test was concluded. If not, the test was continued until 10 min of steady state was achieved or to 45 min of data collection, whichever occurred first. Analyses indicated that RMR estimates from the first 5 min of calorimetry were significantly higher than the remainder of the measures. However, RMR estimates from indirect calorimetry segments meeting steady state criteria or from calorimetry extended beyond 30 min did not significantly differ from those at 30 min (26). Therefore, we used the 5–30-min segment of calorimetry measurements for our calculation of RMR and the mean of the 2 RMR measures in our estimation of TEE.

Activity assessment

We measured AREE with a uniaxial accelerometer (Caltrac; Muscle Dynamics Fitness Network, Torrance, CA), a 7-cm² unit that is worn at the waist and that measures vertical accelerations of the body’s center of gravity. When movement occurs, a cantilevered beam in the monitor bends and emits a current proportional to the force acting on it. A computer in the monitor plots an acceleration curve and uses the area under the curve for the estimation of activity (27).

We instructed participants to maintain their usual activities while wearing the monitor during waking hours for 3 consecutive days, including one weekend day. We input female sex and each woman’s height, weight, and age into her unit at visit 1. The participants cleared the screen to zero and attached the unit to their waistbands or belts under the right arm on arising each day. As needed, they used weightlifting and pedal modes for stationary, isotonic activities and for bicycling or rowing, which switched the accelerometer settings to prediction equations that more closely estimated the energy expended for these activities. Before retiring, participants recorded the readings in the “CALS USED ACTM” screen reflecting AREE only. Water activities precluded the wearing of the unit. The energy expenditure for these activities, though rare (n = 9 of 278 d), was added to the day’s monitor total (28). We included data if the accelerometer was worn for ≥ 22 waking h/d collectively and if the participant followed protocol instructions. We used the mean of the accelerometer measurements for 2–3 d for our estimate of TEE. Each participant completed the Physical Activity Scale for the Elderly (PASE) questionnaire before visit 1 (29), and scores were calculated for comparison with accelerometer data. The PASE questionnaire assesses the frequency and duration of low-, moderate-, and
high-intensity activities over the previous week, and it has been used in studies of physical activity levels in older adults (30). The scale gives a total score that can range from 0 to >400, with higher scores representing higher activity levels.

**Subject characteristics**

*Anthropometry*  
We measured height, weight, and waist and hip circumferences at each clinic visit. We estimated body composition by using urinary creatinine concentrations from two 24-h urine collections 1 wk apart. Participants documented collection times and completeness, stored their urine at ≤4.4°C in a refrigerator or cooler at home, and submitted collections at each clinic visit. Urinary creatinine concentrations were determined with a kinetic modification of the Jaffé alkaline picrate-reaction procedure using a Cobas Mira Plus Analyzer (Roche Diagnostics, Brandenburg, NJ) according to the manufacturer's instructions. The interassay CVs for low, medium, and high urine quality-control pool levels were 1.2%, 1.6%, and 1.6%, respectively. Daily creatinine excretions were excluded if they were <780 mg/d or if urine was collected for <23 h or >25 h (31). We were unable to estimate body composition for 7 women because of urine collections that were incomplete, as calculated by Welles's formula for older adults (32). The percentage of body fat (%BF) was derived from the fat-free mass (FFM) (weight – FFM = kg fat; kg fat/weight × 100% = %BF).

*Psychosocial measures*  
We administered the Crowne-Marlowe Social Desirability Scale during the second visit (33). This is a 33-item, true-false questionnaire that measures a person's tendency to provide the most socially desirable answer regardless of the truth. The tool discriminates between high total scores, representing a strong tendency to choose the socially desirable answer, and low total scores, representing a weaker tendency to do so, at a level of \( P = 0.05 \) or better. The scale has an internal consistency coefficient of 0.88 (Cronbach's alpha) (14, 33).

We used the Stunkard-Sorensen silhouettes to assess the participants' perceptions of body size (34). This instrument shows 9 silhouettes from thin to large, which are given a score of 1 to 9, respectively. Participants were asked to identify the silhouettes that best represented their perceived body size, the desired body size, and the healthiest body size. Dissatisfaction with body size was calculated as perceived body size minus desired body size. We also calculated the difference between current body size and the body size perceived to be the healthiest to determine its influence on energy reporting.

*Other measures*  
We collected data on age, race, education, marital status, and household income from a self-administered questionnaire.

**Statistical analyses**  
We present descriptive data (\( \bar{x} \pm SD \)) for our measures of energy expenditure, the energy intake reported on the FFQ (FFQ energy), and energy underreporting (FFQ energy – TEE), as well as Pearson correlation coefficients for multiple measures. To provide a general assessment of our factorial approach to estimating TEE, we present associations with participant characteristics that should influence TEE [age, body mass index (BMI; in kg/m²), FFM, %BF] and with characteristics that should not influence TEE [waist-to-hip ratio (WHR) and social desirability]. We used multiple linear regression to model the effects of participant characteristics on energy underreporting. The dependent variables were ratio measures of energy reporting. Specifically, we calculated a ratio of FFQ energy divided by TEE for our primary measure of energy reporting error. For analyses, these ratio variables were naturally log transformed to yield approximate normality, and the back-transformed means for FFQ energy divided by TEE are given in tables in the book by Fleiss (35). Independent variables were sociodemographic characteristics (age, income, and education), adiposity-related measures (BMI, FFM, %BF, and WHR), and psychosocial factors (social desirability, perceived body size, dissatisfaction with body size, and disparity from healthy size). To ensure the stability of our estimates, we divided our independent variables into tertiles, except in cases where meaningful categories existed, as follows. Age was divided into decades (50–59, 60–69, and 70–79 y), and education was grouped by level of attainment (high school or some college, bachelor’s degree, and advanced degree). BMI was categorized by using cutoffs from a 1998 consensus conference of the Expert Panel on the Identification, Evaluation, and Treatment of Overweight in Adults: <25, 25–29.9, and ≥30 (36). A WHR > 0.80 is associated with increased health risks in women, and 0.80 was used as an upper cutoff (37).

For purposes of comparison, we also examined the ratio of FFQ energy to RMR. Linear regression models using demographic characteristics and adiposity-related measures were controlled for age, education, and household income. Models using psychosocial factors were controlled for age, education, household income, and BMI. We used a linear contrast to test for trends across ordered categories. Pearson correlation coefficients were calculated for FFQ energy reported and TEE measured, and partial correlations were calculated as for the regression models. The \( z \) scores were used to test for differences between coefficients and for trends (38). PASE score tertiles were compared with average accelerometer measures. Statistical analyses used SAS version 6.12 (SAS Institute, Inc, Cary, NC). Trends were assigned significance if \( P < 0.05 \).

**RESULTS**  
This sample was 94% white, and 60% of the subjects were married or living as married. Thirty-two percent held college degrees, and 69% of those providing income information (\( n = 95 \)) had household incomes of ≥$35 000. Participants had a mean (±SD) BMI of 26.4 ± 4.8, %BF estimates of 30 ± 10%, FFM of 47.8 ± 4.7, and WHR of 0.78 ± 0.07.

The mean values and within-subject Pearson correlation coefficients for our measures of energy expenditure, FFQ energy, and energy underreporting are shown in Table 1. Duplicate indirect calorimetry measures of RMR were highly correlated (\( r = 0.92 \)), as were the first and last days of AREE estimates measured on the accelerometer (\( r = 0.70 \)). Analyses of the accelerometer data using paired \( t \) tests found no significant differences in AREE estimates between weekends and weekdays or between consecutive and nonconsecutive days of data collection (data not shown). FFQ energy at visit 1 and at visit 2 was also highly correlated (\( r = 0.79 \)), and a two-tailed \( t \) test found no significant differences between the 2 measures.

The distribution of estimates of energy-reporting errors identified in this group of 102 healthy postmenopausal women is shown.
TABLE 1  
Measures of energy expenditure and self-reported energy intake in healthy postmenopausal women aged 50–79 y 1

<table>
<thead>
<tr>
<th>Measure</th>
<th>Energy intake 4 (kJ [kcal])</th>
<th>Correlation coefficient (r) 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resting metabolic rate</td>
<td>4904 ± 736 (1172 ± 176)</td>
<td>0.92</td>
</tr>
<tr>
<td>Activity-related energy expenditure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 1 (n = 96)</td>
<td>1946 ± 1121 (465 ± 268)</td>
<td>0.70 6</td>
</tr>
<tr>
<td>Day 2 (n = 92)</td>
<td>1966 ± 971 (470 ± 232)</td>
<td></td>
</tr>
<tr>
<td>Day 3 (n = 90)</td>
<td>2084 ± 1155 (498 ± 276)</td>
<td></td>
</tr>
<tr>
<td>Total energy expenditure</td>
<td>7669 ± 1494 (1833 ± 357)</td>
<td></td>
</tr>
<tr>
<td>Self-reported energy intake</td>
<td>6581 ± 2335 (1573 ± 558)</td>
<td>0.79</td>
</tr>
<tr>
<td>Underreported energy intake</td>
<td>−1105 ± 2452 (−264 ± 586)</td>
<td></td>
</tr>
</tbody>
</table>

1 n = 102 except where indicated otherwise.
2 ± SD.
3 Within-subject duplicate-measure Pearson correlation coefficient.
4 Two-visit average of resting metabolic rate measured between 5 and 30 min of indirect calorimetry.
5 Measured with an accelerometer.
6 Correlation between the first and last days of activity data collection.
7 Estimated from a 2-visit average of resting metabolic rate measured between 5 and 30 min of indirect calorimetry + 2–3-d average of activity-related energy expenditure assessed by accelerometer, with 10% adjustment for thermic effect of food.
8 Average of energy reported on 2 Women’s Health Initiative food-frequency questionnaires (FFQs).
9 FFQ energy — total energy expenditure.

in Figure 1. Thirty-six percent of our sample (n = 37) had underreported energy estimates of ≥ 2400 kJ (574 kcal) according to a comparison of FFQ and TEE estimates. The median for energy underreporting estimates in this sample was 20.8%, ranging from 80% underreporting to 140% overreporting (data not shown).

Correlation coefficients among various measures of energy expenditure, FFQ energy, and energy underreporting are shown in Table 2. TEE was highly correlated with RMR (r = 0.73) and AREE (r = 0.85), as expected, given that TEE is derived from these measures. However, AREE correlated with RMR only at r = 0.26. FFQ energy was only modestly correlated with the objective measures of energy expenditure, but strongly correlated with energy underreporting. That is, as FFQ energy intake increased, energy underreporting decreased (r = 0.81).

To informally assess the validity of our factorial approach to estimating TEE, we present associations of TEE with participant characteristics that should influence TEE (age, BMI, FFM, and %BF) and those that should not influence TEE (WHR and social desirability) (Table 3). As expected, TEE significantly decreased with age (P for trend = 0.0005) and increased with BMI (P for trend = 0.0001), FFM (P for trend = 0.0001), and %BF (P for trend = 0.054). No significant differences were evident by WHR or social desirability. Other characteristics having no significant effect on TEE were education, annual income, and all measures of perceived body size (data not shown).

Data on energy underreporting (FFQ energy divided by TEE) by age, education, household income, and adiposity are given in Table 4. A ratio of 1 would indicate a lack of systematic under–or overreporting for the group of women, whereas a ratio < 1 would indicate underreporting. There was a suggestion of greater underreporting (eg, smaller FFQ energy-to-TEE ratios) with younger age (P for trend = 0.07). Among the adiposity-related factors, BMI was not associated with underreporting of energy intake. The use of other BMI categories such as tertiles or those given in the Surgeon General’s Report on Nutrition and Health (39) also showed no associations with underreporting. Underreporting may increase with increasing FFM (P for trend = 0.11), although this trend was not significant. There were no significant differences in reporting error between participants who were married or living as married and all others (data not shown).

Partial correlation coefficients further illustrating how FFQ energy compares with measured TEE according to participant characteristics are also given in Table 4. There were significantly higher correlation coefficients between FFQ energy and TEE for women with less education (P for trend = 0.04), higher BMI (P for trend = 0.06), and greater FFM (P for trend = 0.03).

Because the accelerometer measure of AREE is novel and vulnerable to bias (participants were aware that we were monitoring their activity levels), we also examined the ratio of FFQ energy to RMR for purposes of comparison (data not shown). Overall, the associations of participant characteristics with ratios of FFQ energy to RMR were similar to those observed for ratios of FFQ energy to TEE, with P values differing only slightly. Specifically, for this outcome, the negative association of educational attainment and underreporting was suggestive (P for trend = 0.06), whereas trends in age (P for trend = 0.17) and FFM (P for trend = 0.20) were not. PASE scores were positively correlated with average accelerometer measures (r = 0.25, P < 0.01). PASE score tertiles (ie, low, medium, and high activity levels) corresponded to accelerometer means of 1723 kJ (412 kcal), 2029 kJ (485 kcal), and 2142 kJ (512 kcal), respectively, with a significant test for trend (P < 0.01).

Data on energy underreporting by psychosocial factors are given in Table 5. Women with high social desirability scores tended to be more likely to underreport energy intakes (ie, have lower FFQ energy-to-TEE ratios) than women with lower scores (P for trend = 0.09) (33). Women perceiving they were thin according to the Stunkard-Sorensen silhouettes (34) tended to underreport their energy intake more than women perceiving themselves as heavy (P for trend = 0.11). Again, similar trends
resulted when ratios of FFQ energy to RMR were used, with a significant association between underreporting energy and social desirability (P for trend = 0.04) (data not shown). A suggestive trend in energy underreporting was seen for dissatisfaction with perceived body size (perceived size − desired size) by use of the FFQ energy–RMR outcome variable (P for trend = 0.09); that is, women closer to their desired size underreported energy more.

Partial correlation coefficients further illustrating how reported FFQ energy compares with measured TEE according to psychosocial characteristics are given in Table 5. There was a higher correlation coefficient between FFQ energy and TEE for women with higher social desirability (P for trend = 0.03), as well as a suggestive association for higher dissatisfaction with perceived body size (P for trend = 0.10).

**DISCUSSION**

These data indicate that healthy postmenopausal women underreported their energy intake by ≈21% on the WHI FFQ. This degree of underreporting is consistent with that seen in doubly labeled water method studies (1, 40–43). Overall, the most important finding is that energy underreporting appeared to vary with participant characteristics, and evidence suggested more underreporting among women who were younger or had high social desirability scores. Johansson et al (2) reported an increase in energy underreporting with increasing age in a younger group (age 44–9 y). However, age differences among populations studied (47, 48), limited sample sizes, or differences in the measures of energy expenditure and energy intake (ie, the FFQ) (49).

We found no association between BMI and energy underreporting such as was found in numerous other studies (3, 4, 6–12). An interesting point regarding the association of adiposity with underreporting is that there is no plausible biologic reason that excess body fat would, in and of itself, cause women to underreport energy intake. Therefore, the various measures of body size and adiposity must be serving as surrogates for psychosocial characteristics that result in underreporting energy, such as poor awareness of intake or portion sizes, deliberate underreporting, and subconscious biasing toward intake that is perceived to be appropriate. We did observe several suggestive associations of psychosocial factors with underreporting. Postmenopausal women

 **TABLE 2** Pearson correlation coefficients between measures of energy expenditure and self-reported energy intake in 102 healthy postmenopausal women

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Activity-related energy expenditure</th>
<th>Total energy expenditure</th>
<th>FFQ energy</th>
<th>Underreported energy intake</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resting metabolic rate</td>
<td>0.26</td>
<td>0.73</td>
<td>0.26</td>
<td>−0.20</td>
</tr>
<tr>
<td>Activity-related energy expenditure</td>
<td>—</td>
<td>—</td>
<td>0.24</td>
<td>−0.38</td>
</tr>
<tr>
<td>Total energy expenditure</td>
<td>—</td>
<td>—</td>
<td>0.24</td>
<td>−0.38</td>
</tr>
<tr>
<td>FFQ energy</td>
<td>—</td>
<td>—</td>
<td>0.24</td>
<td>−0.38</td>
</tr>
</tbody>
</table>

1Correlation coefficients >0.18 are significant, P = 0.05.
2Measured with an accelerometer.
3Estimated from a 2-visit average of resting metabolic rate measured between 5 and 30 min of indirect calorimetry plus 2–3-d average of activity-related energy expenditure assessed by accelerometer, with 10% adjustment for thermic effect of food.
4Two-visit mean energy intake reported on the Women’s Health Initiative food-frequency questionnaire (FFQ).
5Total energy expenditure — total energy expenditure.
6Two-visit average of resting metabolic rate measured between 5 and 30 min of indirect calorimetry.

**TABLE 3** Total energy expenditure by subject characteristics in 102 healthy postmenopausal women aged 50–79

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Unadjusted</th>
<th>Adjusted†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50–59 (n = 47)</td>
<td>8263 (1975)</td>
<td>8192 (1958)</td>
</tr>
<tr>
<td>60–69 (n = 31)</td>
<td>7326 (1751)</td>
<td>7376 (1763)</td>
</tr>
<tr>
<td>70–79 (n = 17)</td>
<td>6523 (1559)</td>
<td>6686 (1598)</td>
</tr>
<tr>
<td>P for trend</td>
<td>0.0001</td>
<td>0.0005</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25 (n = 43)</td>
<td>7238 (1730)</td>
<td>7146 (1708)</td>
</tr>
<tr>
<td>25–29.9 (n = 32)</td>
<td>7439 (1778)</td>
<td>7163 (1712)</td>
</tr>
<tr>
<td>≥30 (n = 20)</td>
<td>8874 (2121)</td>
<td>8761 (2094)</td>
</tr>
<tr>
<td>P for trend</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Fat-free mass (kg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤46 (n = 27)</td>
<td>6774 (1619)</td>
<td>6657 (1591)</td>
</tr>
<tr>
<td>&gt;46 to &lt;49.3 (n = 33)</td>
<td>7728 (1847)</td>
<td>7565 (1808)</td>
</tr>
<tr>
<td>≥49.3 (n = 28)</td>
<td>8535 (2040)</td>
<td>8226 (1966)</td>
</tr>
<tr>
<td>P for trend</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Percentage body fat (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;26 (n = 28)</td>
<td>7481 (1788)</td>
<td>7222 (1726)</td>
</tr>
<tr>
<td>26 to &lt;36 (n = 31)</td>
<td>7368 (1761)</td>
<td>7130 (1704)</td>
</tr>
<tr>
<td>≥36 (n = 29)</td>
<td>8130 (1943)</td>
<td>7987 (1909)</td>
</tr>
<tr>
<td>P for trend</td>
<td>0.084</td>
<td>0.054</td>
</tr>
<tr>
<td>Waist-to-hip ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤0.73 (n = 32)</td>
<td>7401 (1769)</td>
<td>7268 (1737)</td>
</tr>
<tr>
<td>0.74 to 0.80 (n = 36)</td>
<td>7699 (1840)</td>
<td>7376 (1763)</td>
</tr>
<tr>
<td>&gt;0.80 (n = 27)</td>
<td>7870 (1881)</td>
<td>7653 (1829)</td>
</tr>
<tr>
<td>P for trend</td>
<td>0.22</td>
<td>0.31</td>
</tr>
<tr>
<td>Social desirability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low, ≤13 (n = 28)</td>
<td>7958 (1902)</td>
<td>7749 (1852)</td>
</tr>
<tr>
<td>Moderate, &gt;13 to &lt;20 (n = 35)</td>
<td>7481 (1788)</td>
<td>7602 (1817)</td>
</tr>
<tr>
<td>High, ≥20 (n = 31)</td>
<td>7611 (1819)</td>
<td>7770 (1857)</td>
</tr>
<tr>
<td>P for trend</td>
<td>0.36</td>
<td>0.96</td>
</tr>
</tbody>
</table>

1Total energy expenditure (resting metabolic rate + activity-related energy expenditure) with 10% adjustment for thermic effect of food; sample sizes vary because of missing data.
2Age was adjusted for education and income. BMI, fat-free mass, percentage body fat, and waist-to-hip ratio were adjusted for age, education, and income. Social desirability was adjusted for age, education, income, and BMI.
3Crowne-Marlowe Social Desirability Scale (33).
themselves to be thin underreported their energy intake more than
ies (14, 16). We found that postmenopausal women perceiving
trend = 0.09); this finding is consistent with findings in other stud-
underdesirable manner, underreported energy intake more than women with lower scores (P for
trend = 0.09). These findings support the idea that body image, independent of measured BMI, is an important predictor of energy-reporting errors.

This study has many limitations. In particular, the modest sample size limited our ability to detect many associations at conventional significance levels. The generalizability of our results is limited because we included only postmenopausal women from the United States. Although the correlation coefficient for reported FFQ energy and measured TEE (r = 0.24) was similar to the findings of others who used different tools in different populations (all: r = 0.22, P = 0.20) (13, 50), our measure of TEE was likely limited by ≥3 features of the AREE measurement. First, although we specifically instructed women to refrain from strenuous activity for 48 h before the indirect calorimetry measurements,
20 women (19.6%) recorded accelerometer data between 36 and 48 h beforehand. Second, the direction of the measurement error inherent in the accelerometer is unclear. AREE may be underestimated because the accelerometer measures movement in only one plane, does not account for fidgeting or anxiety-related energy expenditures, and is unable to discern variations in grade (51). Third, these participants did not wear the monitor during sleeping hours. One study that compared physical activity estimates determined with this accelerometer to those determined with indirect expenditures, and is unable to discern variations in grade (51). Others described this accelerometer as unable to adequately discriminate between running speeds of 5–8 mph (53). Finally, although physical activity is theoretically an objective measure, participants may have altered their usual physical activity levels in response to wearing an accelerometer (54). RMR alone accounted for 60–70% of TEE (quartiles 1 and 3), and, when we examined the associations of participant characteristics with the ratio of FFQ energy to RMR (which excludes the accelerometer component), trends were almost identical to those seen for the ratio of FFQ energy to TEE, which indicates that errors in the accelerometer measures cannot entirely explain our findings. Nonetheless, one could postulate that errors in this measure of physical activity were associated with participant characteristics (eg, age and BMI) and could influence our results.

The ability to establish relative ranking of dietary energy intake from low to high was affected by participant characteristics, as evidenced by the correlation coefficients between FFQ energy and TEE. For example, dietary energy intake was ranked with more precision in relation to the objective TEE marker among women with lower levels of education, higher income, and greater BMI, social desirability, or dissatisfaction with perceived body size. These observations are limited by small sample sizes, and we know of no similar data in the literature with which to compare our findings.

This study supports the hypothesis that types of errors in dietary energy self-reporting vary by participant characteristics. These errors, when associated with dietary exposures or confounding factors of interest, could result in incorrect inferences regarding diet-disease relations. Studies relying on self-reported dietary intake to assess exposure to dietary components should include objective measures of energy intake (energy expenditure) in at least a subset of participants to identify sources of systematic error within the particular population studied and in the specific self-reporting tool used. The objective measure of energy expenditure in a subset of the study cohort should follow the self-reported intake. These data will allow an assessment of systematic bias for the self-reporting tool in the context of the study population and can be used to calibrate the self-reported data as an integral aspect of diet and disease association analyses. Future objective-measures substudies are needed that have the following characteristics: 1) larger sample sizes, 2) more-diverse populations, 3) additional biomarkers that could assist in determining whether underreporting varies by macronutrient source, and 4) a wider array of psychosocial measures to elucidate the observation that adiposity is associated with underreporting. Identification of characteristics associated with energy-reporting errors can lay the groundwork for the development of statistical methods equipped to adjust for person-specific systematic error in dietary self-reporting.

We express our sincerest thanks to Maggie T Grate and Tamsen Johnston for their assistance with participant visits and materials.

REFERENCES
20. Arciero PJ, Goran MI, Gardner AM, Ades PA, Tyzbir RS, Poehlman ET.
38. Kaczkowski CH, Jones PJH, Feng J, Bayley HS. Four-day multimedia diet records underestimate energy needs in middle-aged and elderly women as determined by doubly labeled water. J Nutr 2000;130:802–5.