European Journal of Clinical Nutrition (2008) 62, 544–552
 0954-3007/08 $30.00
www.nature.com/ejcn

ORIGINAL ARTICLE

Identifying sources of reporting error using measured food intake

WV Rumpler[1], M Kramer[2], DG Rhodes[3], AJ Moshfegh[3] and DR Paul[1]

[1]*Diet and Human Performance Laboratory, US Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Beltsville, MD, USA;* [2]*Biometrical Consulting, US Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Beltsville, MD, USA and* [3]*Food Surveys Research Group, US Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Beltsville, MD, USA*

Objective: To investigate the magnitude and relative contribution of different sources of measurement errors present in the estimation of food intake via the 24-h recall technique.
Design: We applied variance decomposition methods to the difference between data obtained from the USDA's Automated Multiple Pass Method (AMPM) 24-h recall technique and measured food intake (MFI) from a 16-week cafeteria-style feeding study. The average and the variance of biases, defined as the difference between AMPM and MFI, were analyzed by macronutrient content, subject and nine categories of foods.
Subjects: Twelve healthy, lean men (age, 39 ± 9 year; weight, 79.9 ± 8.3 kg; and BMI, 24.1 ± 1.4 kg/m^2).
Results: Mean food intakes for AMPM and MFI were not significantly different (no overall bias), but within-subject differences for energy (EI), protein, fat and carbohydrate intakes were 14, 18, 23 and 15% of daily intake, respectively. Mass (incorrect portion size) and deletion (subject did not report foods eaten) errors were each responsible for about one-third of the total error. Vegetables constituted 8% of EI but represented > 25% of the error across macronutrients, whereas grains that contributed 32% of EI contributed only 12% of the error across macronutrients.
Conclusions: Although the major sources of reporting error were mass and deletion errors, individual subjects differed widely in the magnitude and types of errors they made.
European Journal of Clinical Nutrition (2008) **62,** 544–552; doi:10.1038/sj.ejcn.1602742; published online 11 April 2007

Keywords: energy intake; bias; diet assessment; nutrition assessment; questionnaires

Introduction

To determine the relationship between diet and health, nutritionists rely extensively on self-reported retrospective questionnaires (Food Frequency Questionnaires and 24 h recalls) and food record techniques (recording of food items by a subject over the course of several days). Unfortunately, reporting error is known to accompany these methods (Kipnis *et al.*, 2001, 2003), which can obscure causal relationships between food intake and disease. Reporting error has been attributed to a number of both physical (e.g. body mass index (BMI)) and psychological (e.g. dietary restraint) characteristics, gender and socioeconomic status (Hill and Davies, 2001; Trabulsi and Schoeller, 2001).

The source(s) of food intake reporting error are poorly understood because true intake necessary to compare with questionnaire results is typically unknown. In the absence of knowing true intake, doubly labeled water and urinary nitrogen collection techniques have been proposed as estimates of energy and protein intake, respectively (Bingham, 1994; Trabulsi and Schoeller, 2001; Subar *et al.*, 2003). The limitation of these techniques is that they can only identify misreporting of total energy and protein intakes, and not the misreporting of other macronutrients and micronutrients and/or specific foods. Also, these methodologies provide data averaged over several days and therefore do not allow examination of misreporting on a single day, let alone for a single meal. In addition, they are expensive to incorporate in larger studies and require substantial effort by subjects and investigators.

Correspondence: Dr WV Rumpler, Diet and Human Performance Laboratory, USDA, Agricultural Research Service, 307B Center Road, Beltsville, MD 20705, USA.
E-mail: william.rumpler@ars.usda.gov
Received 27 March 2006; revised 29 January 2007; accepted 29 January 2007; published online 11 April 2007

To understand reporting error on an item-by-item basis, true food intake must be known. Although collecting these data would be prohibitive on a large scale, it is manageable on a small scale. We used a unique data set where all food items actually consumed (measured food intake (MFI)) was known and could be compared to reported values from the USDA's 24-h Automated Multiple Pass Method (AMPM). We used a variance decomposition methodology to determine what percent of the reporting error (difference between true and reported values) was due to, for example forgetting to report an item consumed or misreporting the item's size. To understand better reporting error, decompositions were done by macronutrient category, food group and subject.

Methods

Subjects
Twelve healthy, non-smoking men (eight Caucasians, two African-Americans and two Asians) with an average ($\pm$s.d.) age of 39 ± 9 year, weight of 79.9 ± 8.3 kg and BMI of 24.1 ± 1.4 kg/m^2 participated in this study. Subjects were weight stable and were not using medication known to affect food intake, appetite or water balance. The Johns Hopkins University Bloomberg School of Public Health Committee on Human Research approved the study protocol. Subjects provided written informed consent and received a medical evaluation by a physician, including measurement of blood pressure and analysis of fasting blood and urine samples to screen for metabolic disease.

USDA's automated multiple pass 24-h method
The Beltsville Human Nutrition Research Center (BHNRC) 24-h dietary recall method uses a five-step computer-assisted interview, the AMPM. The AMPM guides the interviewer through the expanded five-pass recall. First, respondents are asked to list, without interruption, all foods and beverages consumed in 24 h on the day before the interview ('quick list'). Second, respondents are asked about forgotten foods and answer a series of questions probing for any forgotten foods from nine categories (non-alcoholic beverages, alcoholic beverages, sweets, savory snacks, fruits, vegetables, cheeses, breads and rolls and any other foods). Third, they are asked about the time and name of eating occasion for each food reported. Fourth, a series of standardized questions probe for detailed information about each food reported and the mass of the food eaten ('detail cycle'). Additional information is elicited about where the food or most ingredients were obtained from and if each eating occasion was at home. Eating occasions and times between occasions are also reviewed. Fifth, respondents are then asked whether anything else was consumed. The design of the interview includes standardized questions and possible response options for the many types of foods available in the United States, with each response option programmed to be followed by the next appropriate question.

The USDA's AMPM was conducted on two occasions over the telephone with each subject, during weeks 14–16 of the *ad libitum* study described below, on different weekdays of two consecutive weeks. The AMPM data were then paired with the MFI data for these days. Energy and macronutrient composition were determined by labels when available and, when not available, there were consultation with the USDA Nutrient Database for Standard Reference Release 15 (US Department of Agriculture ARS, 2002).

Ad libitum *feedings*. The methods for the *ad libitum* feeding have been published previously (Paul *et al*., 2005). Briefly, each day for 16 weeks, subjects consumed only meals that were provided by the Human Studies Facility at the BHNRC. Subjects chose foods *ad libitum* from the menus and were instructed to consume any part or all of a food item and return the remaining portion to be weighed. BHNRC staff who came into contact with the subjects provided no guidance as to the quantities and/or types of food items chosen. Beverages other than those provided by the BHNRC, including those containing alcohol, were allowed but not provided. Subjects reported the brand and exact quantity of beverages in a log that was submitted daily to one of the investigators. As we did not know the true intake for these items, we accepted the values as without measurement error. However, the number of these items was small (12 items) and thereby did not represent a substantial contribution to the total.

Menus
Food items offered in the morning were presented in a cafeteria-style setting as three different rotating menus, each lasting 7 days. Some food items remained on all three menus and were accessible in the evening. Additional food items offered for dinner rotated every 15 days. Three main factors were important in designing menus and selecting food items: (1) allow detection of macronutrient selection effect by offering a wide range of carbohydrate-, fat- and protein-rich foods, (2) provide a variety of commonly eaten foods typical of those many Americans eat, and (3) efficiently utilize the research kitchen staff and resources available. More than 300 food items were used to develop menus for this study and over 100 food items were offered on any given day. Whenever possible, specific requests for food items were incorporated into the menus. To ensure that subjects were allowed free selection of foods that contained a wide range of macronutrient compositions, the menus always contained foods classified as high-carbohydrate, -fat and -protein. The large number of food items offered and long rotation period for the menus was essential for removing cues that might help subjects recall food items chosen (e.g. 'cooked carrots are always offered on Tuesdays').

Statistical analysis and decomposition of variance

We define reporting error (sometimes referred to as measurement error) as the difference between reported (AMPM) and actual MFI. This quantity is calculated on a per item basis. The reporting error for each item was classified, manually by the first author, into one of four mutually exclusive categories: mass, misclassification, addition and deletion. Mass differences, calculated to the nearest gram, were defined as the difference between AMPM and MFI in the weight of the item and, therefore, was a continuous variable. Misclassification differences were those due to inaccurate description of the food item (such as baked chicken versus fried chicken) or to differences in the composition of what was consumed versus the composition in the database for a given item (i.e. item A had a 'true' fat content of 8 g versus the database content of 6 g). This latter discrepancy occurred in a few number of cases. Additions were foods reported but not actually consumed, wheras deletions were foods not reported but consumed.

There are two components of reporting error (bias) that interested us, the average and the variance. We define the estimated bias (henceforth, bias) as the mean difference between AMPM and MFI. The bias describes whether certain food categories tend to be under- or over-reported. Irrespective of whether bias exists, items could vary in the accuracy they were reported; individual subjects could also differ in their accuracy. The variance (average squared difference between AMPM and MFI for the item) is a measure of how accurately food items are reported. For example, the average reported value for a food item might be very close to the average measured value, but individual portions may not be reported accurately. In this case, the variance would be large but the bias small.

Using standard variance decomposition techniques (Searle *et al*., 1992), the total variance of the reporting error for an item was decomposed into the four categories described above plus the six covariances for pairs of categories (e.g. the covariance between mass and misclassification). The covariance is a measure of whether, within a subject or food group, one type of reporting error tends to be accompanied by another. For each of the macronutrient groups (energy, carbohydrate, protein, fat), the total (TOT) variance of the reporting error is decomposed into four component parts: mass (E_M), misclassification (E_C), addition (E_A), and deletion (E_D). If a subject substituted one clearly different item for another, the discrepancy was considered to be an addition and a deletion rather than a misclassification.

There are three steps to calculate the quantities discussed in Results, namely, (1) assigning a measurement error value to every individual food item in the study, further partitioned by macronutrient category, (2) grouping these values by subject (or food group for one analysis), and calculating variance and covariance estimates for the four types of reporting errors, and (3) calculating the total variance, as the sum of the variances and covariances calculated in (2). These steps are explained in detail below.

For each food item, if there was a difference between reported and measured values, we assigned one of the four reporting errors (mass, misclassification, addition or deletion). In very few cases, it was difficult to determine which type of reporting error occurred. Mass error, E_M, is $E_M = (\text{Mass}_{\text{AMPM}} - \text{Mass}_{\text{MFI}}) \times \text{Composition}_{\text{MFI}}$, where $\text{Composition}_{\text{MFI}}$ is a conversion factor (e.g. converting the mass into energy or grams of fat, depending on the macronutrient category being investigated). Misclassification error is $E_C = (\text{Mass}_{\text{AMPM}} \times \text{Composition}_{\text{AMPM}}) - (\text{Mass}_{\text{MFI}} \times \text{Composition}_{\text{MFI}})$. Using the two conversion factors allows for the misclassified food to have, for example, a different percentage of fat than what was reported. Addition error, E_A, is simply $\text{Mass}_{\text{AMPM}}$, that is, the amount the subject reported to have eaten but did not eat and deletion error, E_D, is Mass_{MFI}, the measured amount eaten but not reported. In cases where no difference existed between measured and reported, all errors were set to zero.

These values were calculated for each of the five macronutrients: amount (g), energy (J), carbohydrate (g), protein (g), and fat (g). Thus, for the *i*th item, we have either no reporting error or error assigned to one of the four types of reporting error, and we then calculated that error for each of the five macronutrients.

The next step is to group the items by subject (or food type) and calculate variances and covariances (calculated separately for each of the five macronutrients). If we group by subject for energy, the average reporting errors in energy for each of the 12 subjects, $R_{jk} = \frac{1}{n}\sum_{i=1}^{n}(E_{ij})$, where *i* indexes the *n* food items associated with subject *k* and *j* specifies the type of reporting error (only one of these can be nonzero for a given item; all four are zero if there is no reporting error for the given item). The averages for other macronutrients are calculated similarly. We can then calculate variances and covariances in the standard way, $\text{Var}(R_j) = \frac{1}{12}\sum_{k=1}^{12}(R_{jk} - \bar{R}_j)^2$ y and $\text{Cov}(R_j, R_m) = \frac{1}{12}\sum_{k=1}^{12}[(R_{jk} - \bar{R}_j)(R_{mk} - \bar{R}_m)]$, where *j* and *m* refer to two different types of reporting error.

This analysis was also carried out for each food group separately, for example, only items in the grains food group (see below) were used. We also calculated the variances and covariances for each subject separately by first grouping by food group, that is separately for each subject calculate R_{jk}, but here *k* indexes food groups rather than subjects, and then use the same variance and covariance formulas given above.

The total variance that we calculate for each of the decompositions is $\sum_{j=1}^{4}\text{Var}(R_j) + 2\sum_{j \neq m}^{6}\text{Cov}(R_j, R_m)$. We use this value as a denominator when calculating the percent of the total variance attributable to each of the four kinds of reporting errors.

We used the VARCOMP procedure from SAS/STAT software (Version 8, SAS Institute Inc., Cary, NC, USA) for calculating the variance decompositions for food categories and subjects. The nine categories of foods (Figure 1), we used were described by Krebs-Smith *et al*. (2000): grain products (grains; G), fruits and juices (F), vegetables (V), milk-yogurt-cheese (dairy; D), meat-fish-poultry (meats; M), beverages (B),

Grain products (G)
- Yeast bread
- Crackers
- Muffins/biscuits
- Pancakes/waffles/French toast
- Cooked cereal
- Ready-to-eat cereal
- Rice, other cooked grains/mixtures
- Pasta/pasta mixture
- Pizza
- Doughnuts/sweet rolls
- Cookies/brownies
- Cake/pie
- Chips/popcorn/pretzels

Fruits (F)
- Fruit juice
- Fruit

Vegetables (V)
- White potatoes
- Lettuce, other greens
- Other vegetables

Milk, yogurt, cheese (D)
- Milk on cereal
- Milk in coffee or tea
- Milk as a beverage
- Cheese
- Yogurt

Meat, fish, or poultry (M)
- Eggs, excluding mixtures
- Meat, fish, poultry or egg sandwich/ mixture

Beverages (B)
- Beer
- Wine
- Coffee, tea
- Soft drinks, regular
- Soft drinks, diet

Sweets (S)
- Candy
- Sweet spreads/syrups
- Artificial sweeteners

Fats (L)
- Fat-type spreads
- Cream/creamers, not whipped
- Cream/creamers, liquid
- Creamers, powdered
- Dressings, not mayo-type
- Mayo-type dressing

Other (O)
- Soups
- Nuts/seeds, butters
- Frozen dairy desserts

Condiments

Figure 1 Categories of foods. Adapted from Krebs-Smith *et al.* (2000).

sweets (S), fats and oils (F) and other (O). The coefficients of variation (CV) for reporting error we report are calculated as the square root of the average reporting error variance for items in the category divided by the mean of the items (using measured values and calculated over subjects). We tested the difference between reported and measured means (bias) using paired *t*-tests (pairing within subjects).

Results

All 12 subjects completed two AMPM interviews (24 AMPM–MFI pairs). Of the 538 food items actually consumed, only 510 food items were reported (20 additions and 48 deletions). These food items represented 162 of the 350 foods offered. There were nine alcohol-containing beverages or food items consumed with one addition and no deletion. Alcohol intake was included in total energy intake but, owing to the small number of items, no separate analysis of alcohol intake was conducted.

Macronutrient intake

Although mean differences (averaged over individuals) between AMPM and MFI intakes were not significant (Table 1), there were substantial differences among individuals (Figure 2). The absolute difference (sign ignored) between reported and measured energy intake within person averaged 1.56 MJ/day (14% of daily intake, ranging from 1.20 to 3.50 MJ/day). The absolute difference (sign ignored) in reporting error within person for protein, fat and carbohydrate averaged 18, 23 and 15% of daily intake, respectively.

Table 2 presents the results from decomposing the variance of reporting error into the four categories of error and their covariance. Mass (portion size) and deletion errors both represented about a third of the variance associated with total energy, carbohydrate and fat intake. Approximately half of the error for protein intake was associated with mass and an additional third was due to misclassification. Addition errors contributed 26% of the total for carbohydrate, but only 17% for energy, 10% for fat and 6% for protein. All covariances were small (based on our definition of error categories, a different description of the error model would result in different covariances).

For each of the food categories, not grouped by subject, the number of reporting errors ($\geqslant$1 g/item), the percent of daily energy intake and the percentage of the total error in reported macronutrient intake contributed by the category are presented in Table 3. Errors were present in 507 of 538 items consumed. Vegetables constituted 8% of the total energy intake, but represented the largest portion of the error across macronutrients. Beverages also constituted a small percentage of the daily energy intake (5%), but had a disproportionate contribution to the error in energy intake (17%). Grains, which constituted the largest portion of total

Table 1 Reported and measured total food and beverage, and macronutrient intakes for men consuming *ad libitum* diets

	Total food and beverage, g/day (s.d.)	*Energy, MJ/day (s.d.)*	*Carbohydrate, g/day (s.d.)*	*Protein, g/day (s.d.)*	*Fat, g/day (s.d.)*
Overall means					
Measured	2728 (573)	10.9 (2.2)	342 (88)	92.9 (20.8)	103.3 (37.1)
Reported	2724 (698)	10.7 (2.3)	343 (98)	99.0 (25.8)	89.7 (24.3)
Intake day 1					
Measured	2850 (572)	10.7 (1.9)	341 (70)	95.1 (22.8)	96.4 (30.8)
Reported	2853 (725)	10.6 (2.1)	345 (85)	104.0 (22.4)	83.6 (19.4)
Intake day 2					
Measured	2606 (573)	11.1 (2.5)	343 (107)	90.7 (19.3)	110.3 (42.7)
Reported	2595 (675)	10.8 (2.6)	340 (113)	94.0 (28.9)	95.9 (27.8)

No differences were significant (paired *t*-test).
Number of observations (*n*) per value is equal to 12.
Measured = a total food and beverage and macronutrient intakes measured during *ad libitum* food intake averaged across subjects.
Reported = total food and beverage and macronutrient intakes measured by 24-hr recall (AMPM) averaged across subjects.
Day 1 = first day of comparison between reported and measured total food and beverage and macronutrient intakes averaged across subjects.
Day 2 = second day of comparison between reported and measured total food and beverage and macronutrient intakes averaged across subjects.

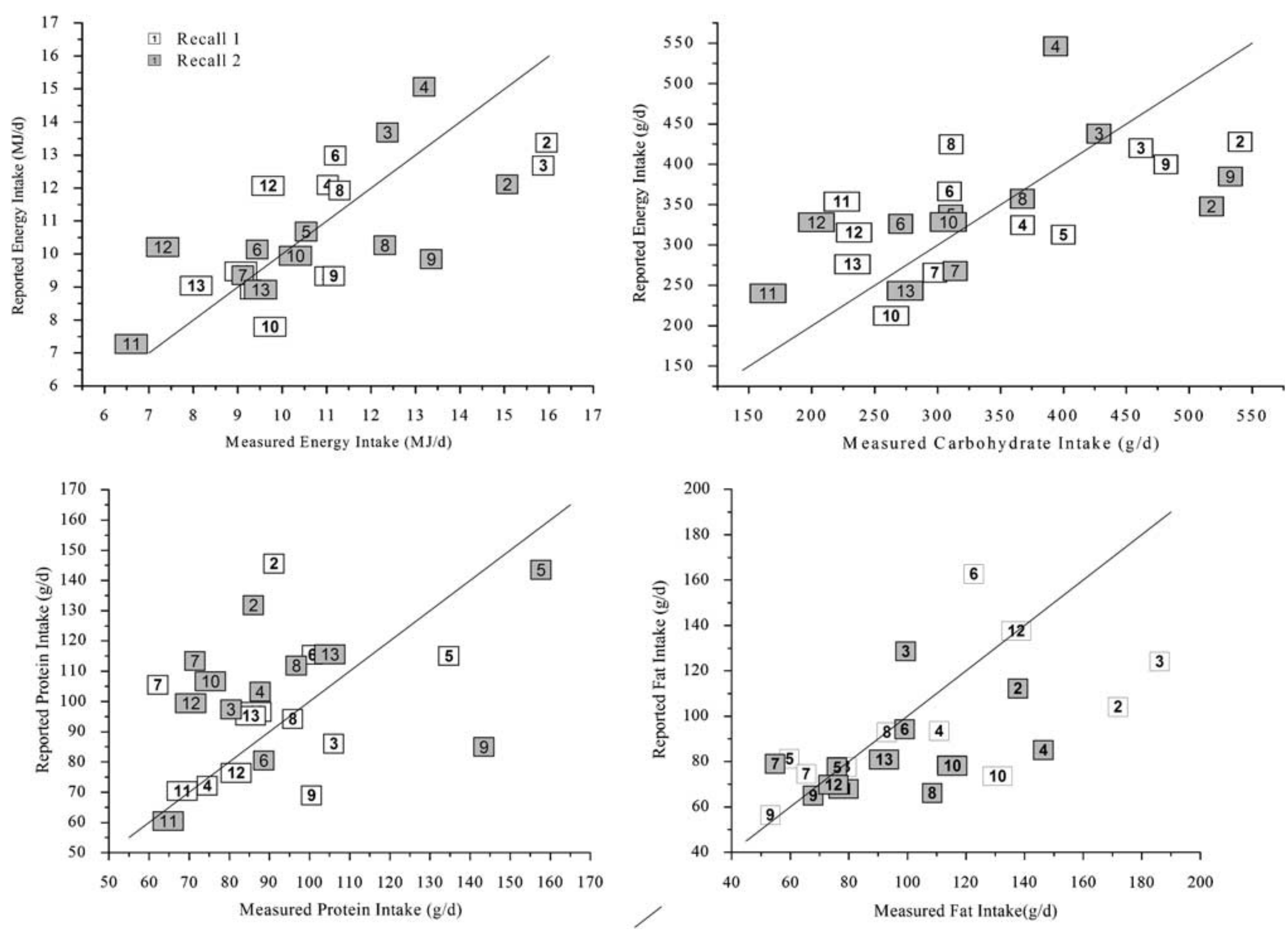

Figure 2 Reported and measured energy, carbohydrate, protein and fat intakes on 2 days separated by at least 1 week of men (identified by the numbers 1–12 in boxes) consuming *ad libitum* diets.

Table 2 Decomposition of the variance of the difference between reported and measured food for men consuming *ad libitum* diets

Variances	*Energy*		*Carbohydrate*		*Protein*		*Fat*	
	(j/item)²	*Estimate % of total*	*(g/item)²*	*Estimate % of total*	*(g/item)²*	*Estimate % of total*	*(g/item)²*	*Estimate % of total*
Total ($n = 538$)	248		0.324		0.022		0.048	
Mass	83	34	0.105	33	0.011	49	0.016	33
Misclassification	30	12	0.029	9	0.006	29	0.013	26
Addition	43	17	0.084	26	0.001	6	0.005	10
Deletion	84	34	0.097	30	0.004	19	0.014	30
Covariances	*(J/item)*		*(J/item)*		*(J/item)*		*(J/item)*	
Mass–misclassification	2.24		0.002		*		*	
Mass–addition	0.13		*		*		*	
Mass–deletion	−0.31		*		*		*	
Misclassification–addition	−0.37		*		*		*	
Misclassification–deletion	0.85		*		*		*	
Addition–deletion	1.50		0.002		*		*	

*|Covariance| < 0.001.
Measured = total food and beverage and macronutrient intakes measured during *ad libitum* food intake.
Reported = total food and beverage and macronutrient intakes measured by 24-h recall (AMPM).
Mass = difference between reported and measured intakes due to incorrect estimation of the mass of an item.
Misclassification = difference between reported and measured intakes due to inaccurate description of an item.
Addition = difference between reported and measured intakes due to reporting of an item that was not actually consumed.
Deletion = difference between reported and measured intakes due to the failure to recall an item that was actually consumed.

Table 3 Number of items, percent of daily energy intake and percent of error in reported macronutrient intake, represented by food categories for men consuming *ad libitum* diets

Category	*No. of food items*	*Items*	*No. of errors*	*Energy*	*Percent error*		
		% (s.d.)			*Carbohydrate*	*Protein*	*Fat*
Beverages	47	5.2 (12.1)	42	17	11	12	2
Dairy	56	9.4 (8.0)	56	9	26	9	7
Fruits and juices	78	13.3 (7.0)	75	10	9	10	5
Grains	107	32.2 (12.9)	94	12	10	14	8
Fats and oils	35	6.0 (6.2)	35	9	3	12	6
Meats	44	12.7 (7.9)	44	3	6	3	2
Other	32	8.9 (8.2)	26	7	9	5	10
Sweets	22	3.9 (4.7)	19	4	4	1	0
Vegetables	117	8.3 (5.5)	116	28	23	33	59
Beverages	47	5.2 (12.1)	42	17	11	12	2

Items = food items consumed per category of food group (independent of subject).
No. of food items = number of food items consumed.
% = percentage of daily energy intake attributed to food group.
No. of errors = number of items misreported.
Percent error = percent error in energy and macronutrient intakes.

energy intake (32%), contributed only 12% of the error in energy intake and a similar contribution to the error across macronutrients. Dairy contributed the largest portion of the error for carbohydrate (26%), but contributed to the error in proportion to its contribution to total intake (<10%) for the other macronutrients.

Coefficients of variation for the reporting errors per item within food category are presented in Table 4. The largest CV in the reporting of individual items occurred in beverage for energy (78%), dairy for carbohydrate (105%), fats and oils for protein (77%) and vegetables for fat (257%). CVs across macronutrients were consistently low for grains. Beverages, which represented only 5% of total energy intake, had high CVs for energy, carbohydrate and protein, but the second lowest CV for fat.

The average bias and the decomposition of the reporting error variance into mass, misclassification, addition and deletion (and the sum of their covariances) by individual food category (within subject) and subject (within food category) are presented for energy intake in Tables 5 and 6. There was great heterogeneity among food groups and subjects in the magnitude and distribution of reporting

Table 4 Coefficient of variation (CV) of the difference between reported and measured food intake by category of foods for men consuming *ad libitum* diets

Category	n	*Energy*	*Carbohydrate*	*Protein*	*Fat*
Beverages	47	78	59	63	28
Dairy	56	31	105	36	61
Fruits and juices	78	26	27	28	37
Grains	107	25	25	33	45
Fats and oils	35	51	20	77	82
Meats	44	14	31	17	20
Other	32	50	83	41	188
Sweets	22	44	47	13	13
Vegetables	117	47	46	62	257

Table 5 Estimated average bias and variance decomposition, across subjects, of reported vs measured energy intake for men consuming *ad libitum* diets

Category	*Bias*		*% of total variance*				*Covariances*
	n	*J/item*	*Mass*	*Misclassification*	*Addition*	*Deletion*	
Total	538	−158.8	34	12	17	34	3
Beverages	47	−18.1	60	11	23	5	1
Dairy	56	−4.3	55	9	0	30	7
Fruits and juices	78	−24.6	27	0	46	26	1
Grains	107	−52.7	26	2	11	64	−2
Fats and oils	35	−95.6	75	3	0	18	4
Meats	44	−120.6	41	8	53	0	−2
Other	32	94.8	23	42	2	27	6
Sweets	22	77.4	99	1	0	0	0
Vegetables	117	−15.1	20	40	4	18	18

Bias = difference between measured and reported energy intake on a per item basis (expressed in joules).
Mass = difference between reported and measured intakes due to incorrect estimation of the mass of an item.
Misclassification = difference between reported and measured intakes due to inaccurate description of an item.
Addition = difference between reported and measured intakes due to reporting of an item that was not actually consumed.
Deletion = difference between reported and measured intakes due to the failure to recall an item that was actually consumed.
Covariances = sum of the six covariances (between pairs of types of reporting errors).

Table 6 Estimated average bias and variance decomposition, across food groups, of reported vs measured energy intake for men (identified by subject number) consuming *ad libitum* diets

Subject	*Bias*		*% of total variance*				*Covariances*
	n	*J/item*	*Mass*	*Misclassification*	*Addition*	*Deletion*	
1	44	42.3	16	2	78	0	4
2	48	−30.6	34	14	19	17	16
3	45	−70.4	52	3	0	58	−13
4	43	0.8	63	10	5	5	17
5	56	−86.4	6	16	0	70	8
6	39	54.9	90	11	0	3	−3
7	53	58.6	56	26	21	6	−8
8	48	−49.2	48	18	0	32	1
9	48	−59.3	26	1	49	21	2
10	35	−109.2	55	19	0	36	−10
11	38	16.9	55	9	10	19	7
12	41	24.2	36	16	12	34	2
Mean		−17.3	44.8	12.1	16.2	25.1	1.9

Bias = difference between reported and measured energy intake on a per subject basis (expressed in joules).
Mass = difference between reported and measured intakes due to incorrect estimation of the mass of an item.
Misclassification = difference between reported and measured intakes due to inaccurate description of an item.
Addition = difference between reported and measured intakes due to reporting of an item that was not actually consumed.
Deletion = difference between reported and measured intakes due to the failure to recall an item that was actually consumed.
Covariances = sum of the six covariances (between pairs of types of reporting errors).

error. Half of the subjects, and eight of the nine food categories were, on average, underreported (bias <0). Mass represented the greatest portion of the error in energy intake for six of the nine food categories and for nine of the 12 subjects. Large addition errors occurred with fruits and juices, meats and for subjects 1 and 9. Large deletion errors occurred with grains and for subjects 3 and 5. Whereas misclassification errors represented a major portion of the variance for vegetables and other food categories, it was no more than one fourth for any individual subject.

Conclusions

The results of this investigation found that the average bias was small for group estimates of energy and macronutrient intakes when administering the AMPM in a group of lean male subjects. However, within-subject estimates of energy and macronutrient intake based on AMPM are much less accurate.

In a study also using the AMPM (although a different version from the one used in this study) in 261 weight stable men, a bias of -11% in energy intake, as estimated from energy expenditure (assumes expenditure = intake in weight stable individuals), was detected (Subar *et al*., 2003). The bias in energy intake with 24 h recalls observed in Subar *et al*. appears to be consistent with the general consensus that energy intake is typically underreported when using criteria such as doubly labeled water (Hill and Davies, 2001; Trabulsi and Schoeller, 2001). However, two investigations that compared the AMPM to a single-day, laboratory-controlled feeding regimen also failed to detect a bias in energy intake in men (Conway *et al*., 2004), despite biases in normal and overweight women (Conway *et al*., 2003). The lack of a bias from the current study vs. the clear underreporting from studies such as Subar *et al*. (2003) may be related to the large variation in day-to-day food intake (Rumpler *et al*., 2006) and the inability to represent adequately this variation with a few 24-h recalls. However, it may also be due to subject-specific factors such as age and/or body composition (Novotny *et al*., 2003). The subjects from the current study were 39.9 years (28–50 years) and 24 kg/m^2, whereas the subjects from Subar *et al*. (2003) were all above 40 years (63% were over 50 years) and 78% were $\geqslant$25 kg/m^2 (30% were $\geqslant$30 kg/m^2). It is likely that these factors have a role in reporting bias (Hill and Davies, 2001). Note that, despite little overall reporting error bias for these subjects, the considerable within-subject error (and substantial bias for some subjects) indicates that the bias is imprecisely estimated.

The two major sources of discrepancy between reported and measured were poor estimation of mass (portion size) and failure to report foods that were consumed (deletion). Poor estimation of portion size has been long recognized (Lansky and Brownell, 1982; Weber *et al*., 2001), yet it presents a significant problem for estimating energy intake (Harnack *et al*., 2004). Reports of greater than 20% over- and underestimation are consistent with our observations (Godwin *et al*., 2004). In many cases, there were practical explanations for why these and other types of errors occurred. For example, some of the vegetable misclassification errors were due to subjects reporting that the baked and refried beans were regular fat, rather than the low fat versions that were actually on the menus. For meats, these additions and deletions were generally the result of reporting something consumed on a previous day. The additions for fruit and juices were generally the result of reporting something consumed regularly but not on the day of the AMPM (i.e., orange juice daily for breakfast).

The results of this study indicate that adjusting for reporting error based on 'known' discrepancies in a single macronutrient or total energy intake may introduce substantial bias in the estimate of the other macronutrients (Bellach *et al*., 1998; Paul *et al*., 2005). Reported food intake is routinely adjusted by normalizing to an independent estimate of energy intake (Black *et al*., 1991; Black and Cole, 2000; Barnard *et al*., 2002). On the basis of our results, a more detailed understanding of reporting error by food category is required. For example, grains accounted for 32% energy intake, whereas accounting for only 12% of the total misreporting. This is contrasted by beverages, which constituted only 5% of the total energy intake but contributed to 17% of total error. Thus, to accurately adjust reported intake to reflect actual intake, a detailed knowledge of the type and extent of misreporting by food category is required.

This study was not intended for use in generalizing the results of these 12 men to a wider population. It is likely that our subject population (well-educated, lean men) are one of the least susceptible to measurement error, and biases and variances may be larger if we had selected a population that are more likely to misreport (e.g., overweight, high dietary restraint, women). However, this analysis provides some insight into the interpretation of reported food intake. It is worth noting that even these subjects, who should be able to report their food intakes accurately, often had large errors, and the types of errors may differ among subjects.

We suggest three types of information that might be useful to reduce the differences between reported and actual food intake. First, identify covariates (such as physical characteristics or memory) that are related to accurate food intake reporting. The consistent biases of some subjects suggest that reporting error can be reduced if additional characteristics influencing responses were known. Second, food groups might be reorganized in a way to better identify and adjust for errors. For example, rather than to organize by food type, it might be better to perform a *post hoc* organization by the type of error made. Thus, instead of grouping granola bars (usually eaten as a snack, where deletion would be a common problem) with granola (usually eaten for breakfast, where portion size estimations might be a common problem), they could be grouped with foods producing similar kinds of errors. These new groupings would then be easier to adjust as similar errors would be made for all items in a group. Third, identify error-prone food groups or items, and modify questionnaires to reduce reporting error accordingly. Questions about items known to produce a certain kind of error should be reviewed to determine if modifications can be made to reduce the risk of making that kind of error. Minimizing reporting error should reduce the need for and magnitude of overall corrections, which may overcorrect some food groups and under correct others (Paul *et al.*, 2005).

References

Barnard J, Tapsell L, Davies P, Brenninger V, Storlien L (2002). Relationship between high energy expenditure and variation in dietary intake with reporting accuracy on 7 day food records and diet histories in a group of healthy adult volunteers. *Eur J Clin Nutr* **56**, 358–367.

Bellach B, Kohlmeier L (1998). Energy adjustment does not control for differential bias in nutritional epidemiology. *J Clin Epidemiol* **51**, 395–398.

Bingham S (1994). The use of 24-h urine samples and energy expenditure to validate dietary assessments. *Am J Clin Nutr* **59**, 227S–231S.

Black A, Cole T (2000). Within- and between-subject variation in energy expenditure measured by the doubly-labelled water technique: implications for validating reported dietary energy intake. *Eur J Clin Nutr* **54**, 386–394.

Black A, Goldberg G, Jebb S, Livingstone M, Cole T, Prentice A (1991). Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. *Eur J Clin Nutr* **45**, 583–599.

Conway JM, Ingwersen LA, Moshfegh AJ (2004). Accuracy of dietary recall using the USDA five-step multiple-pass method in men: an observational validation study. *J Am Diet Assoc* **104**, 595–603.

Conway JM, Ingwersen LA, Vinyard BT, Moshfegh AJ (2003). Effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women. *Am J Clin Nutr* **77**, 1171–1178.

Godwin SL, Chambers E, Cleveland L (2004). Accuracy of reporting dietary intake using various portion-size aids in-person and via telephone. *Journal of the American Dietetic Association* **104**, 585–594.

Harnack L, Steffen L, Arnett DK, Gao S, Luepker RV (2004). Accuracy of estimation of large food portions. *J Am Diet Assoc* **104**, 804–806.

Hill RJ, Davies PS (2001). The validity of self-reported energy intake as determined using the doubly labelled water technique. *Br J Nutr* **85**, 415–430.

Kipnis V, Midthune D, Freedman LS, Bingham S, Schatzkin A, Subar A *et al.* (2001). Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am J Epidemiol* **153**, 394–403.

Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP *et al.* (2003). Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol* **158**, 14–21; discussion 22–16.

Krebs-Smith SM, Graubard BI, Kahle LL, Subar AF, Cleveland LE, Ballard-Barbash R (2000). Low energy reporters vs others: a comparison of reported food intakes. *Eur J Clin Nutr* **54**, 281–287.

Lansky D, Brownell KD (1982). Estimates of food quantity and calories: errors in self-report among obese patients. *Am J Clin Nutr* **35**, 727–732.

Novotny JA, Rumpler WV, Riddick H, Hebert JR, Rhodes D, Judd JT *et al.* (2003). Personality characteristics as predictors of under-reporting of energy intake on 24-h dietary recall interviews. *J Am Diet Assoc* **103**, 1146–1151.

Paul DR, Rhodes DG, Kramer M, Baer DJ, Rumpler WV (2005). Validation of a food frequency questionnaire by direct measurement of habitual ad libitum food intake. *Am J Epidemiol* **162**, 806–814.

Rumpler WV, Kramer M, Rhodes DG, Paul DR (2006). The impact of the covert manipulation of macronutrient intake on energy intake and the variability in daily food intake in nonobese men. *Int J O (2005)* **30**, 774–781.

Searle S, Casella G, McCulloch CE (1992). *Variance Components*. John Wiley and Sons: New York.

Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S *et al.* (2003). Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* **158**, 1–13.

Trabulsi J, Schoeller DA (2001). Evaluation of dietary assessment instruments against doubly labeled water, a biomarker of habitual energy intake. *Am J Physiol Endocrinol Metab* **281**, E891–E899.

US Department of Agriculture ARS (2002). USDA Nutrient Database for Standard Reference, Release 15. In:U.S. Government Printing Office: Washington, DC. Nutrient Data Laboratory Home Page, http://www.nal.usda.gov/fnic/foodcomp.

Weber J, Reid P, Greaves K, DeLany J, Standford V, Going S *et al.* (2001). Validity of self-reported energy intake in lean and obese young women, using two nutrient databases, compared with total energy expenditure by doubly labeled water. *Eur J Clin Nutr* **55**, 940–950.