Published in final edited form as:
Ann Epidemiol. 2009 August ; 19(8): 553–559. doi:10.1016/j.annepidem.2009.04.010.

Number of 24-Hour Diet Recalls Needed to Estimate Energy Intake

Yunsheng MA, MD, PhD, **Barbara C. Olendzki, RD, MPH**, **Sherry L. Pagoto, PhD**, **Thomas G. Hurley, MS**, **Robert P. Magner, DC, MPH**, **Ira S. Ockene, MD**, **Kristin L. Schneider, PhD**, **Philip A. Merriam, MSPH**, and **James R. Hébert, MSPH, SCD**
Division of Preventive and Behavioral Medicine (Y.M, B.C.O., S.L.P., R.P.M., K.L.S., P.A.M.) and the Division of Cardiovascular Medicine (I.S.O.), Department of Medicine, University of Massachusetts Medical School, Worcester; and the Cancer Prevention and Control Program and Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia (T.G.H., J.R.H.)

Abstract

Purpose—Twenty-four-hour diet recall interviews (24HRs) are used to assess diet and to validate other diet assessment instruments. Therefore it is important to know how many 24HRs are required to describe an individual's intake.

Method—Seventy-nine middle-aged white women completed seven 24HRs over a 14-day period, during which energy expenditure (EE) was determined by the doubly labeled water method (DLW). Mean daily intakes were compared to DLW-derived EE using paired *t* tests. Linear mixed models were used to evaluate the effect of call sequence and day of the week on 24HR-derived energy intake while adjusting for education, relative body weight, social desirability, and an interaction between call sequence and social desirability.

Results—Mean EE from DLW was 2115 kcal/day. Adjusted 24HR-derived energy intake was lowest at call 1 (1501 kcal/day); significantly higher energy intake was observed at calls 2 and 3 (2246 and 2315 kcal/day, respectively). Energy intake on Friday was significantly lower than on Sunday. Averaging energy intake from the first two calls better approximated true energy expenditure than did the first call, and averaging the first three calls further improved the estimate ($p = 0.02$ for both comparisons). Additional calls did not improve estimation.

Conclusions—Energy intake is underreported on the first 24HR. Three 24HRs appear optimal for estimating energy intake.

Keywords

Doubly Labeled Water; Energy Expenditure; Energy Intake; Mental Recall; Nutrition Assessment; Underreporting

Introduction

Diet plays a role in the etiology and prevention of many chronic diseases such as coronary heart disease, diabetes, and cancer (1–3). Accurate estimation of dietary intake is essential for

Address correspondence to: Dr Yunsheng Ma, Division of Preventive and Behavioral Medicine, Department of Medicine, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655. Tel: (508) 856-1008. Fax: (508) 856-2022. Yunsheng.Ma@umassmed.edu.

Study concept and design: Y.M., J.R.H., B.C.O., I.S.O.; statistical analysis: R.P.M., Y.M.; drafting of the manuscript: Y.M.; critical revision of the manuscript for important intellectual content: J.R.H., B.C.O., S.L.P., T.G.H., I.S.O., R.P.M., K.S., and P.A.M.

None of the authors had any conflicts of interest.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

assessing the effect of diet on health, as well as the effectiveness of weight loss and other lifestyle interventions of which diet is an important part. Because total energy intake is a determinant of both the nutrient content of the diet and individual nutrient requirements, it is crucial to obtain accurate estimates of intake.

For a variety of practical reasons, nutritional data in most epidemiologic studies are obtained from self-report or interviewer-administered dietary assessment methods. One such method, the 24-hour dietary recall interview (24HR), is commonly used both as an intensive method of assessment for intervention studies and as a comparison method for validation/calibration studies of structured assessments such as food frequency questionnaires (4,5). Some large epidemiologic studies may collect 24HR data as part of a validation substudy (6); more rarely, they may collect one or two days of 24HR data as part of the general assessment protocol (7, 8).

As diets vary considerably from day to day (9–13), the ability of a single 24HR to provide an accurate estimate of long-term energy intake is limited. However, the degree to which energy intake estimates improve with increasing number of recalls is not known. Several studies have attempted to answer this question, with varying results. Researchers have suggested that 3, 4, 5, or 7 days are necessary to adequately estimate energy intake (14–17). In addition, the number of recalls required may differ by age, ethnicity, or other characteristics due to variability in eating habits (11,12,16,17). Because neither an objective criterion, such as the doubly labeled water (DLW) method, was employed nor were subjects followed up under highly controlled laboratory conditions, true energy intake is unknown in these studies. Therefore the intra-individual/inter-individual variance ratio is often used to estimate the number of recalls needed (18).

The objectives of the present study were to 1) evaluate effect of call sequence on estimates of energy intake and 2) calculate the number of 24HRs needed to accurately estimate the group mean energy intake. Energy expenditure (EE) was determined by the DLW method, which is the gold standard for assessing energy intake. For individuals in weight balance (i.e., neither gaining nor losing body weight), EE equals energy intake (19); therefore EE provides an accurate estimate of true energy intake.

Materials and Methods

Subjects

Eighty-one women were recruited from a university campus and surrounding communities between June and October 1997. Women were eligible if they 1) were free of major medical conditions; 2) were not taking steroid-based medications, such as asthma inhalers or prednisone; and 3) had a body weight less than 91 kg (200 lb). Women were excluded if they 1) were on a special diet to lose or gain weight; 2) would not consent to maintaining their current physical activity and dietary habits during the study; or 3) were not available to be reached by telephone on any day during the study period. Premenopausal subjects were scheduled to enter the study the week after they completed menstruation.

Participants were followed for 14 days; during which 7 days of 24HRs were performed. True energy intake was assessed by DLW (19). The study protocol was approved by the Institutional Review Board of the University of Massachusetts Medical School.

Total EE from DLW

A mixed $^{2}H_2{}^{18}O$ dose containing 1.5 g/kg body weight of $H_2{}^{18}O$ (10 atom % excess) and 0.05 g/kg body weight of $^{2}H_2O$ (>99 atom % excess) was given orally at the beginning of the clinic visit, followed by a 100-mL tap water rinse. Spot urine samples (pre-dose, 1-, 7-, and 14-day post-dose) were collected and stored at −80°C until analysis. Samples were analyzed by isotope ratio mass spectrometry at Metabolic Solutions, Inc. (Nashua, NH) using Europa instrumentation (Europa Scientific Ltd, Crewe, UK). Total EE was calculated from carbon dioxide production using an estimate of respiratory quotient (20). Detailed DLW administration and total EE calculation are described elsewhere (19).

Assessment of Diet

Dietary data were collected from participants on 7 of the 14 days of their study enrollment; the intent was that each day of the week be assessed once for each participant. The 24HR multiple-pass method probes for complete food descriptions, detailed food preparation methods, and diverse portion size descriptions (21). Prior to administration of the 24HR, each subject received two-dimensional food models with depiction of foods and serving sizes in order to assist in reporting portion sizes of food intake. Subjects were instructed to have these available for the calls. The dietitians entered dietary data directly into an interactive nutrition analysis software program, the Nutrition Data System for Research (NDS) software (Version 2.91, Nutrition Coordinating Center, Minneapolis, MN). Each 24HR took approximately 30 minutes. All missing foods (foods not found in the NDS database) were resolved through matching similar nutrient content to existing foods in the database, or through the Nutrition Coordinating Center, which provides new nutrient and ingredient formulas for submitted foods. Total energy intake from each 24HR was calculated using the NDS software.

Psychosocial, Body Weight and Other Assessments

Social desirability was measured by the Marlowe-Crowne Social Desirability Scale (22). Increasing scores are associated with increased social desirability, that is, the perceived need to defend oneself in a situation perceived to be a test. Social approval was measured by the Martin-Larsen Approval Motivation Scale (23), for which increasing scores are associated with increased social approval (i.e., the perceived need to seek approval in a situation perceived to be a test). Both were completed at baseline. Anthropometric measures were measured on day 0 and day 14 and included body weight, height, waist circumference (smallest circumference between lower rib and the iliac crest), abdomen circumference (umbilicus), and hip circumference (largest protrusion). Fat-free mass, total body water, and body fat were calculated from DLW-derived data (19).

At baseline, leisure-time physical activity was determined by self-administered questionnaire using assessment and scoring methods developed by the Behavioral Risk Factor Surveillance System (24). Women with reported activity for 20 or more minutes three or more times per week were termed as having “regular activity”, women with less than 20 minutes or less than three times per week were termed as having “irregular physical activity,” and women with no reported leisure-time physical activity were termed as “sedentary.”

Demographic information, such as age, race, and educational level, also was collected at baseline.

Statistical Analyses

Participants' characteristics and nutritional variables were summarized by using mean and standard deviation for continuous variables and number (percentage) for categorical variables. Linear mixed models were used to evaluate the unadjusted effect of call sequence and day of the week on self-reported energy intake. Subject identification was fit as a random effect. We then adjusted these estimates for covariates including education and social desirability score, previously found to bias self-reported dietary intake (19,25–27). Additional variables considered for inclusion in the final analyses were body mass index (BMI = weight [in kilograms]/height [in square meters]) and social approval score, as they also were found to be

NIH-PA Author Manuscript

NIH-PA Author Manuscript

associated with underreporting (28,29). However, social approval score was not significantly associated with energy intake and therefore was not included in the final model. Possible interactions among independent variables were examined; a significant interaction term between social desirability and call sequence was included in the final model.

Unadjusted average energy intake was computed by the number of recalls in increasing sequence order (e.g., an average for calls 1 and 2, then a separate average for calls 1–3, up to a 7-day average); these multi-day averages were compared separately to the estimate of daily EE derived from DLW using a paired *t* test. In addition, a paired *t* test was used to compare each multi-day average estimate of energy intake to the average energy intake for all 7 days of 24HR.

Results

A total of 79 (97.5%) healthy women (mean age = 49.1 years, standard deviation [SD] = 6.8) completed seven 24HRs during the 14-day metabolic period during which total EE was determined by using the DLW method. We excluded data from one woman without DLW data and from another who completed only six 24HRs.

Participants' characteristics are summarized in Table 1. Most of the participants were White (97.5%), well educated (87% with some college or more education), and nonsmoking (89.9%). Nearly half (49.4%) were predominantly sedentary in their leisure-time. Slightly more than half were premenopausal. Their average BMI was 27.0 kg/m^2 (SD = 10.4), total EE from DLW was 2114.8 kcal/day (SD = 405.7), and average energy intake from 7 days of 24HRs was 1825.6 kcal/day (SD = 466.7).

Table 2 presents daily energy intake from 24HR by call sequence and day of the week obtained from the linear mixed models. In both the raw and adjusted results, a call sequence effect was found: daily energy intake derived from the first call was the lowest of any day. Both the raw and adjusted means increased significantly at the second and third calls. The raw mean for the fourth call increased slightly, whereas the adjusted value decreased to a point intermediate between the first and second call. Adjusted daily energy intakes from the fourth to seventh calls and the unadjusted daily energy intakes from the fifth to seventh calls were not significantly different from the first call. Daily energy intake on Friday (adjusted mean = 1746 kcal) was significantly lower than Sunday (adjusted mean = 1906 kcal); adjusted energy intake for the other days of the week was not significantly different from Sunday. We noted that there was a significant interaction between social desirability score and call sequence, with the most pronounced effects observed on the second and third call (amounting to underestimates of 31.5 and 30.1 kcal/point per day, respectively; both $p < 0.05$). The effect was completely attenuated after the third 24HR.

Table 3 presents the difference between estimates of energy intake from the 7-day average from 24HR and DLW with the multi-day averages computed using recalls from days 2–6. The reported *p* values were obtained from paired *t* tests. Averaging the energy intake from the first two calls (mean = 1768.8) provided a significantly better approximation of daily total EE from the DLW compared to a single call (mean = 1672.3) ($p = 0.02$). Averaging the first three calls (mean = 1815.1) provided further significant improvement ($p = 0.02$). Including additional calls did not significantly improve energy estimation. Energy differences between 24HR and 7-day-average 24HR by number of recalls also were compared; similar findings were observed such that two 24HRs provided improvement over a single call, three 24HRs provided further improvement, and more than three calls did not improve energy estimation. It should be noted that although our intention was to have each day represented once per person, there was some imbalance in that Sunday, Wednesday, and Thursday were underrepresented (72, 69, and 53 times, respectively), whereas Monday, Tuesday, Friday, and Saturday were over-represented (80, 109, 84, and 86 times, respectively).

Discussion

Results from the study indicate that call sequence effect should be considered in collecting 24HR-derived dietary data. The first call, regardless of day of the week, was associated with significant underreporting. Two 24HRs provided an increase of 96 kcal/day over a single call. Three 24HRs, which provided improvement of 143 kcal over a single call, produced the best estimate of energy intake. Although three 24HRs have been recommended for better estimation of energy intake (9), our study extends the literature by suggesting that the first 24HR is associated with significant underreporting of energy intake. Additionally, collecting more than three recalls did not significantly improve energy estimation, despite substantially increased cost.

Few studies have examined the effect of call sequence on energy estimation. In a previous study, done in a Latino population, we found that participants' trust in the caller increased with increasing number of 24HRs, with the callers reporting greater ease in subsequent 24HRs (28). The first 24HR took longer than subsequent 24HRs because it included time spent with introductions, social conventions, and discussion of topics that may not be directly related to 24HR (28). It seems as though some degree of comfort/familiarity on the part of the participant is necessary to produce results similar to those obtained from the criterion DLW-derived EE value.

While most studies expend very little effort training subjects to improve their reporting accuracy, underestimation in the first 24HR may be related to portion size estimation and social conventions involved in the interview. Novotny et al. (30) showed that the number of food omissions were higher with the first 24HR than with a subsequent 24HR and suggested that this may be due to a training effect of repeat interviews (30). The finding of significant underreporting of energy intake from the first 24HR has important implications for studies involving 24HR methodologies. Originally, the National Health and Nutrition Examination Survey collected data in its nationally representative sample using one 24HR. The National Health and Nutrition Examination Survey 2001–2002 survey started collecting two 24HRs in a 10% subsample in 2001, and then in all subjects in 2002. Based on our findings, developing methods to adjust energy intake from the first recall might be warranted. Results from the present study argue against using a single recall to estimate individual level energy intake because of the large amount of error, including underestimation of total energy intake.

Because 24HRs are expensive, sampling strategies in randomized clinical trials often limit the number of days of information collected to 2 or 3, with 1 or 2 weekdays and 1 weekend day to capture both energy and nutrient variability of the diet (9). Johnson and colleagues conducted a study to determine the accuracy of energy intake from three 24HRs compared to total energy expenditure estimated from the DLW over a 14-day period under free-living conditions among 24 children between the ages of 4 and 7 years (31). No difference was found between 3-day mean energy intake and total energy expenditure for the group. They concluded that three 24HRs were sufficient to make valid group estimates of energy intake. However, sequence effect was not examined. A recent study by Moshfegh and colleagues from the U.S. Department of Agriculture (USDA) (32) conducted a much larger study (524 participants, aged 30–69 years old) validating the 24HR method against DLW. Each subject was dosed with DLW on the first day of the 2-week study period; three 24HRs were collected during the 2-week period, the first recall was conducted at the first day in person, and subsequent recalls were over the telephone: scheduled 5 to 6 and 10 to 11 days later, respectively. Overall, the subjects underreported energy intake by 11% when compared to DLW-derived EE. The USDA data also point out the need to consider BMI as a confounder in the accuracy of dietary recall. To obtain more robust estimates, we also had data on covariates for adjustment including BMI, education, and social desirability and social approval, response sets previously found to bias self-reported dietary intake (29). While both our study and the USDA study concluded that underreporting is an issue, we had seven 24HRs, and our study affirms that the first 24HR represents underreporting of energy intake. Three 24HRs are required for the best estimate of energy intake. One very important finding from our study is that additional calls did not improve energy estimation.

In this study we found that the second and third administration of the 24HR was associated with social desirability bias. It is interesting that the effect was attenuated after the third day, but overall estimation did not improve. Agreement with DLW-derived energy expenditure was good, especially for adjusted energy intake. Also, anticipating the bias means that it can be controlled statistically.

Because, on average, reported intakes were higher on weekends than during the week, at least one weekend day should be included in order to reduce error in energy intake. Finally, because the first recall represents significant underreporting, an introductory call may be useful before the first 24HR is administered. The introduction should include education for portion size estimation using visual comparison; for example, using a plate or one's hand, paying special attention to educational level and culture. We found this practice has been useful in conducting 24HR in the Latino population (28). Alternatively, if resources are abundant, one could discard the first recall, and average recalls from the second through the fourth 24HRs for best approximation of energy intake.

There are several strengths of our study. First, we had true EE estimated from DLW and a larger sample size than is typically used in studies of DLW. The use of DLW-derived measures of EE enabled us to base our estimate of the number of recalls required on comparisons with an objective measure of intakes. Second, the collection of seven recalls distributed over a variety of different days also enables us to evaluate how days of the week, as well as the numberofcalls, influence the accuracy of reporting. Finally, we were able to adjust for several factors known to affect energy intake including BMI, education, and social desirability bias (19).

The use of a true criterion measure is rare in nutritional epidemiology and calibrating in a highly compliant population represents a good first step in understanding call sequence and other effects. We measured DLW during a 14-day period; the amount of weight (indeed, body compartment mass) change was close to zero (19). Women on special diets to lose or gain weight were excluded from participation. So, the data completion rate was very high and control for extraneous factors was excellent.

Several weaknesses of our study are worth noting. First, participants were middle-aged women with an average BMI less than 30 kg/m^2, thereby limiting our ability to generalize the results to men or obese individuals. We recruited women who had a body weight less than 91 kg (200 lb) for two reasons: 1) to make sure the results were generalizable to the vast majority of women, who weigh less than 90 kg; and 2) to conserve $^2H_2{}^{18}O$, which was (and is) extraordinarily expensive. Second, all participants were highly motivated and willing to undergo an intensive battery of dietary and other assessments. Third, despite our attempt at balancing the days of the week, there was some imbalance. Fourth, because the best estimate obtained from the recalls represented a mean level of underreporting of more than 20% versus DLW measures, we acknowledge that the recalls were unable to estimate energy intakes in absolute terms. Finally, data used for this investigation were collected over 10 years ago; and patterns of dietary intake may have changed since then. Although overall intake may have increased, reporting ability should not have changed. Therefore, even though there is a secular trend toward increased

BMI, there is no good reason to believe that the relationship between eating, weight regulation, and self-reported diet changed materially over the interval.

In conclusion, we found that the first 24HR represents a significant underreporting of energy intake. Two 24HRs provided improvement over a single call. Three 24HRs were sufficient to minimize the mean difference between reported and objectively measured intakes. Additional calls did not improve energy estimation. In addition, we found that, on average, reported intakes were higher on weekends than during the week. Future studies should investigate ways to improve dietary assessments.

Acknowledgments

The project was supported by National Institute of Diabetes and Digestive and Kidney Diseases grant DK 52079-02 to Dr James R. Hébert. The authors thank Ms Susan Druker, the project director, and the women who participated in the study.

References

1. Ma, Y.; Hebert, J.; Ebbeling, C.; Ockene, I. International aspects of coronary heart disease epidemiology. In: Becker, RC.; Alpert, JS., editors. Cardiovascular medicine-practice and management. London: Arnold; 2001.
2. Liu S. Whole-grain foods, dietary fiber, and type 2 diabetes: searching for a kernel of truth. Am J Clin Nutr 2003;77:527–529. [PubMed: 12600838]
3. Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective. American Institute for Cancer Research; Washington (DC): 2007.
4. Hebert JR, Gupta PC, Bhonsle RB, Sinor PN, Mehta H, Mehta FS. Development and testing of a quantitative food frequency questionnaire for use in Gujarat, India. Public Health Nutr 1999;2:39–50. [PubMed: 10452730]
5. Hebert JR, Ockene IS, Hurley TG, Luippold R, Well AD, Harmatz MG. Development and Testing of a Seven Day Dietary Recall. J Clin Epidemiol 1997;50:925–937. [PubMed: 9291878]
6. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122:51–65. [PubMed: 4014201]
7. Murphy SP. Collection and analysis of intake data from the integrated survey. J Nutr 2003;133:585S–589S. [PubMed: 12566508]
8. Dwyer J, Picciano MF, Raiten DJ. Collection of food and dietary supplement intake data: What We Eat in America-NHANES. J Nutr 2003;133:590S–600S. [PubMed: 12566509]
9. Beaton GH, Milner J, McGuire V, Feather TE, Little JA. Source of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr 1983;37:986–995. [PubMed: 6846242]
10. De Castro JM. Weekly rhythms of spontaneous nutrient intake and meal pattern of humans. Physiol Behav 1991;30:729–738. [PubMed: 1775547]
11. Cai H, Shu XO, Hebert JR, Jin F, Yang G, Liu DK, et al. Variation in nutrient intakes among women in Shanghai, China. Eur J Clin Nutr 2004;58:1604–1611. [PubMed: 15199383]
12. Hebert J, Gupta P, Mehta H, Ebbeling C, Bhonsle R, Varghese F. Sources of variability in dietary intake in two distinct regions of rural India: implications for nutrition study design and interpretation. Eur J Clin Nutr 2000;54:479–486. [PubMed: 10878649]
13. Hebert JR, Hurley TG, Chiriboga DE, Barone J. A comparison of selected nutrient intakes derived from three diet assessment methods used in a low-fat maintenance trial. Public Health Nutr 1998;1:207–214. [PubMed: 10933420]
14. Basiotis PP, Welsh SO, Cronin FJ, Kelsay JL, Mertz W. Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. J Nutr 1987;117:1638–1641. [PubMed: 3655942]

15. Nelson M, Black AE, Morris JA, Cole TJ. Between- and within-subject variation in nutrient intake from infancy to old age: estimating the number of days required to rank intakes with desired precision. Am J Clin Nutr 1989;50:155–167. [PubMed: 2750688]
16. Bingham S, McNeil NI, Cummings JH. The diet of individuals: a study of a randomly-chosen cross section of British adults in a Cambridgeshire village. Br J Nutr 1981;45:23–35. [PubMed: 6258625]
17. Wassertheil-Smoller S, Davis BR, Breuer B, Chang CJ, Oberman A, Blaufox MD. Differences in precision of dietary estimates among different population subgroups. Ann Epidemiol 1993;3:619–628. [PubMed: 7921310]
18. Mennen LI, Bertrais S, Galan P, Arnault N, Potier de Couray G, Hercberg S. The use of computerised 24 h dietary recalls in the French SU.VI.MAX Study: number of recalls required. Eur J Clin Nutr 2002;56:659–665. [PubMed: 12080407]
19. Hebert J, Ebbeling C, Matthews C, Ma Y, Clemow L, Hurley T, et al. Systematic errors in middle-aged women's estimates of energy intake: comparing three self-report measures to total energy expenditure from doubly labeled water. Ann Epidemiol 2002;12:577–586. [PubMed: 12495831]
20. Black AE, Prentice AM, Coward WA. Use of food quotients to predict respiratory quotients for the doubly-labelled water method of measuring energy expenditure. Hum Nutr Clin Nutr 1986;40:381–391. [PubMed: 3771290]
21. Schakel SF, Sievert YA, Buzzard IM. Sources of data for developing and maintaining a nutrient database. J Am Diet Assoc 1988;88:1268–1271. [PubMed: 3171020]
22. Crowne D, Marlowe D. A new scale of social desirability independent of psychopathology. J Consult Clin Psychol 1960;24:349–354.
23. Martin HJ. A revised measure of approval motivation and its relationship to social desirability. J Pers Assess 1984;48:508–519. [PubMed: 16367511]
24. Pereira MA, FitzerGerald SJ, Gregg EW, Joswiak ML, Ryan WJ, Suminski RR, et al. A collection of Physical Activity Questionnaires for health-related research. Med Sci Sports Exerc 1997;29:S1–205. [PubMed: 9243481]
25. Hebert JR, Clemow L, Pbert L, Ockene IS, Ockene JK. Social desirability and approval biases in dietary self-report may profoundly compromise the validity of diet-disease studies. Int J Epidemiol 1995;24:389–398. [PubMed: 7635601]
26. Hebert JR, Peterson KE, Hurley TG, Stoddard AM, Cohen N, Field AE, et al. The effect of social desirability trait on self-reported dietary measures among multi-ethnic female health center employees. Ann Epidemiol 2001;11:417–427. [PubMed: 11454501]
27. Hebert JR, Hurley TG, Peterson KE, Resnicow K, Thompson FE, Yaroch AL, et al. Social desirability trait influences on self-reported dietary measures among diverse participants in a multicenter multiple risk factor trial. J Nutr 2008;138:226S–234S. [PubMed: 18156429]
28. Olendzki B, Ma Y, Pagoto S, Hebert J, Griffith J, Merriam P, et al. Underreporting of energy intake and associated factors in a Latino population at risk of developing type 2 diabetes. J Am Diet Assoc 2008;108:1003–1008. [PubMed: 18502234]
29. Hebert JR, Ma Y, Clemow L, Ockene IS, Saperia G, Stanek E, et al. Gender difference in social desirability and social approval bias in dietary self-report. Am J Epidemiol 1997;146:1046–1055. [PubMed: 9420529]
30. Novotny JA, Rumpler WV, Judd JT, Riddick PH, Rhodes D, McDowell M, et al. Diet interviews of subject pairs: how different persons recall eating the same foods. J Am Diet Assoc 2001;101:1189–1193. [PubMed: 11678490]
31. Johnson RK, Driscoll P, Goran MI. Comparison of multiple-pass 24-hour recall estimates of energy intake with total energy expenditure determined by the doubly labeled water method in young children. J Am Diet Assoc 1996;96:1140–1144. [PubMed: 8906138]
32. Moshfegh AJ, Rhodes DG, Baer DJ, Murayi T, Clemens JC, Rumpler WV, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. J Nutr 2008;88:324–332.

Selected Abbreviations and Acronyms

24HR	24-hour dietary recall interview

DLW	doubly labeled water
EE	energy expenditure
SD	standard deviation
USDA	U.S. Department of Agriculture

NIH-PA Author Manuscript
NIH-PA Author Manuscript

TABLE 1

Descriptive statistics, The Energy Study (*N* = 79), Worcester, Massachusetts, June–October 1997

	N	%
Categorical variables		
Marital status		
Single	5	6.3
Married	50	63.3
Living with partner	7	8.9
Divorced	13	16.5
Widowed	4	5.1
Race/ethnicity		
White	77	97.5
Hispanic	2	2.5
Menopausal status		
Premenopausal	43	54.4
Postmenopausal	36	45.6
Education		
High school or less	10	12.7
Some college	35	44.3
Bachelor degree	15	19.0
Graduate school	19	24.0
Employment status		
Full time	49	62.0
Part time	16	20.3
Unemployed	10	12.7
Retired	4	5.1
Occupational classification		
Skill or craft	4	6.2
Scientific work	4	6.2
Service work	3	4.6
Clinical, office or sales	21	32.3
Professional, managerial work	33	50.8
Current smoking status		
Yes	7	10.1
No	72	89.9
Physical activity		
Sedentary	39	49.4
Irregular physical activity	18	22.7
Regular physical activity	22	27.9
Continuous variables	Mean	SD
Age (yr)	49.1	6.8
Anthropometric measures		
Body weight (kg)	69.7	10.4

	N	**%**
Height (cm)	161.1	6.5
BMI (kg/m^2)	27.0	4.0
Lean body mass (kg)*	43.0	5.4
Waist circumference (cm)	84.0	9.6
Abdomen circumference (cm)	95.4	11.8
Hip circumference (cm)	104.3	8.9
Elbow circumference (cm)	7.1	0.6
Body weight (kg) change (day 14-day 0)	0.1	0.8
Total energy expenditure from DLW (kcal/d)	2114.8	405.7
Daily energy intake from 24-hour dietary recal (kcal/d)		
First	1672.3	508.4
7-day average	1825.6	466.7
Social desirability score†	17.4	5.7
Social approval score‡	50.0	8.2

BMI = body mass index; SD = standard deviation.

*Lean body mass was calculated on the basis of double-labeled water, as described in the text.

†Increasing scores are associated with increased social desirability; that is, the perceived need to defend oneself in a situation perceived to be a test. Each of the scale's 33 questions is scores 0 or 1 (depending on a “true” or “false” response; 18 are scored “1” on a “true” response and 15 on a “false” response). Thus social desirability scores fall in the range of 0 to 33. The mean is usually around 17–18 and the Interquartile range is usually around 7, that is, from 15 to 22 (19,25,26,29).

‡Increasing scores are associated with increased social approval, that is, the perceived need to seek approval in a situation perceived to be a test. Each item can take on a value of 1 to 5 (according to the Likert scale responses). Of the 20 questions, 15 items are direct scored (i.e., 1 = 1 to 5 = 5). Five items are reverse scored (i.e., 1 = 5 to 5 = 1). Scores can range from 20 to 100. The mean is usually around 50 and the interquartile range is usually around 12, that is, from 45 to 57 (19,29).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

TABLE 2

Unadjusted and adjusted energy estimates by call sequence and day of the week: Results of linear mixed models, The Energy Study (N = 79), Worcester, Massachusetts, June–October 1997

	N	**Unadjusted energy intake Mean (SE)***	***p* Value**	**Adjusted energy intake Mean (SE)†**	***p* Value**
Call sequence					
1	79	1672.3 (57.2)	Reference	1500.9 (200.9)	Reference
2	79	1865.4 (84.9)	0.02	2246.4 (156.5)	0.007
3	79	1907.7 (84.4)	0.003	2315.3 (139.5)	0.001
4	79	1946.6 (83.3)	0.005	1704.3 (200.5)	0.52
5	79	1853.4 (73.0)	0.04	1667.7 (189.4)	0.56
6	79	1716.6 (73.6)	0.62	1513.1 (193.6)	0.97
7	79	1817.7 (69.5)	0.11	1831.7 (197.3)	0.25
Days of the week					
Sunday	72	1937.0 (73.1)	Reference	1906.1 (83.2)	Reference
Monday	80	1751.0 (76.5)	0.16	1858.0 (79.5)	0.59
Tuesday	109	1721.0 (62.2)	0.01	1813.1 (75.3)	0.27
Wednesday	69	1674.7 (78.2)	0.007	1786.8 (76.9)	0.19
Thursday	53	1887.6 (87.4)	0.55	1915.9 (80.4)	0.99
Friday	84	1922.5 (76.0)	0.04	1746.0 (64.5)	0.04
Saturday	86	1922.8 (81.2)	0.61	1797.5 (73.2)	0.14

SE = standard error.

* Unadjusted means were estimated by using a linear mixed model, which included day of the week or call sequence as independent variables. Subject ID was fit as random effect. Energy intake from each call was the dependent variable.

† Adjusted means were estimated by using a linear mixed model, which included day of the week, call sequence, social desirability score, education, and an interaction between social desirability score and call sequence in independent variables. Subject ID was fit as a random effect. Energy intake from each call was the dependent variable.

TABLE 3

Energy difference between 24-hour recalls and DLW/7-day average by number recalls: The Energy Study (*N* = 79), Worcester, Massachusetts, June–October 1997

	Energy intake Mean (SD)	Energy difference: Recall-DLW Mean (SD)	Energy difference: Recall-7 day avg. Mean (SD)
Energy expenditure from DLW	2114.8 (405.7)		
7-Day 24-hour recalls average	1825.6 (466.7)		
Calls			
1	1672.3 (508.4)	−442.5 (563.6)[*]	−153.4 (393.2)[a]
2	1768.8 (536.5)	−346.0 (576.5)[†]	−56.8 (270.1)[b]
3	1815.1 (575.5)	−299.7 (603.2)[‡]	−10.5 (247.4)[c]
4	1848.0 (529.1)	−266.9 (542.7)[‡]	22.3 (165.3)[c]
5	1849.1 (508.9)	−265.8 (545.4)[‡]	23.4 (139.4)[c]
6	1827.0 (483.5)	−287.9 (531.3)[‡]	−1.3 (78.6)[c]
7	1825.6 (466.7)	−289.2 (512.9)[‡]	–

SD = standard deviation; DLW = doubly labeled water.

[*, †, ‡, a, b, c] Value with the same marker was significantly different by paired *t* test.