American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2018.

Vol. 187, No. 10
DOI: 10.1093/aje/kwy126
Advance Access publication:
June 18, 2018

Practice of Epidemiology

Combining a Food Frequency Questionnaire With 24-Hour Recalls to Increase the Precision of Estimation of Usual Dietary Intakes—Evidence From the Validation Studies Pooling Project

Laurence S. Freedman*, Douglas Midthune, Lenore Arab, Ross L. Prentice, Amy F. Subar, Walter Willett, Marian L. Neuhouser, Lesley F. Tinker, and Victor Kipnis

* Correspondence to Dr. Laurence S. Freedman, Biostatistics Unit, Gertner Institute for Epidemiology, Sheba Medical Center, Tel Hashomer 52621, Israel (e-mail: lsf@actcom.co.il).

Initially submitted December 21, 2017; accepted for publication June 8, 2018.

Improving estimates of individuals' dietary intakes is key to obtaining more reliable evidence for diet-health relationships from nutritional cohort studies. One approach to improvement is combining information from different self-report instruments. Previous work evaluated the gains obtained from combining information from a food frequency questionnaire (FFQ) and multiple 24-hour recalls (24HRs), based on assuming that 24HRs provide unbiased measures of individual intakes. Here we evaluate the same approach of combining instruments but base it on the better assumption that recovery biomarkers provide unbiased measures of individual intakes. Our analysis uses data from the 5 large validation studies included in the Validation Studies Pooling Project: the Observing Protein and Energy Nutrition Study (1999–2000), the Automated Multiple-Pass Method validation study (2002–2004), the Energetics Study (2006–2009), the Nutrition Biomarker Study (2004–2005), and the Nutrition and Physical Activity Assessment Study (2007–2009). The data included intakes of energy, protein, potassium, and sodium. Under a time-varying usual-intake model analysis, the combination of an FFQ with 4 24HRs improved correlations with true intake for predicted protein density, potassium density, and sodium density (range, 0.39–0.61) in comparison with use of a single FFQ (range, 0.34–0.50). Absolute increases in correlation ranged from 0.02 to 0.26, depending on nutrient and sex, with an average increase of 0.14. Based on unbiased recovery biomarker evaluation for these nutrients, we confirm that combining an FFQ with multiple 24HRs modestly improves the accuracy of estimates of individual intakes.

cohort studies; dietary measurement; energy; measurement error; potassium; protein; recovery biomarkers; sodium

Abbreviations: AMPM, Automated Multiple-Pass Method; FFQ, food frequency questionnaire; 24HR, 24-hour recall; NBS, Nutrition Biomarker Study; NPAAS, Nutrition and Physical Activity Assessment Study; OPEN, Observing Protein and Energy Nutrition.

Measurement error in self-reported dietary intakes limits the reliability of results from nutritional cohort studies (1). Approaches to improving the accuracy of dietary intake estimates include reporting intake using new technologies (2, 3), combining self-reports with biomarkers (4), and combining different self-report instruments (5). We focus on the last.

Carroll et al. (5) considered combining information from multiple 24-hour recalls (24HRs) and a food frequency questionnaire (FFQ), using data from the Eating at America's Table Study (6). Their evaluation assumed that 24HRs unbiasedly measure an individual's intake, and they showed that their results were probably robust to departures from this assumption.

Here, we also evaluate combining 24HRs with an FFQ, but under the better-founded assumption that recovery biomarkers unbiasedly measure individual intake (7). We use data from 5 large validation studies with recovery biomarkers as reference instruments, comprising the Validation Studies Pooling Project (8).

Also motivating this work was previous work using the time-varying intake model (9), which showed that correlations between intakes reported using 24HRs and true usual intake were lower

than previously estimated, because of the close proximity in time of biomarker and 24HR assessments in some studies (10). This raised doubts regarding the benefit accruing from adding 24HR assessments to an FFQ. Using recovery biomarker and self-report data, we estimated correlations of intakes estimated from self-report instruments and their combinations with true longer-term usual intakes for energy, protein, potassium, and sodium and their densities. We report that combining instruments led to increased correlation of estimated intakes with true intakes, and below we discuss the implications of these findings.

Determining the optimal combination of 24HR and FFQ data requires knowing the error models of the two instruments, which in turn requires recovery biomarker measurements. Here, we focus on cohort studies having FFQ and 24HR data but no biomarker data. The FFQ and 24HR data would then be combined, as the best option available, under the (usually erroneous) assumption that 24HRs are unbiased. We use recovery biomarkers to provide an unbiased assessment of this method of combining instruments.

METHODS

The Validation Studies Pooling Project

Investigators in 5 large (>200 participants) validation studies using recovery biomarkers agreed to pool their data, aiming, through common analysis, to clarify the nature and magnitude of reporting errors in FFQs and 24HRs (8, 11). The 5 studies in the Validation Studies Pooling Project included diverse populations within the United States. The Observing Protein and Energy Nutrition (OPEN) Study (12) and the Automated Multiple-Pass Method (AMPM) validation study (13) included volunteers aged 40–69 years (OPEN) or 30–69 years (AMPM) residing in Maryland. The Energetics Study included younger white and African-American adults residing in California (14). The Nutrition Biomarker Study (NBS) (15) and the Nutrition and Physical Activity Assessment Study (NPAAS) (16) included postmenopausal women, mostly aged ≥60 years and residing throughout the United States, in the Women's Health Initiative Dietary Modification Trial and the Women's Health Initiative Observational Study, respectively. Further details are provided elsewhere (12–16).

At least 1 FFQ was administered to each participant. This analysis includes only the first administration. The FFQs queried about dietary intake over the past year (OPEN, Energetics, AMPM) or the past 3 months (NBS, NPAAS). One of 3 FFQs was used: the Harvard FFQ (17) (AMPM), the Women's Health Initiative FFQ (18, 19) (NBS, NPAAS), or the Diet History Questionnaire (6) (OPEN, Energetics).

Each study included 2 or more 24HR assessments, administered to all participants in 4 studies and to a 20% subset in the NBS. An interviewer-administered multiple-pass method (4 studies) or a Web-based self-administered 24HR (Energetics) was used. Further details are provided elsewhere (12–16).

Each study measured doubly labeled water for energy intake (20) and collected 24-hour urine samples for measuring nitrogen (21), potassium (22), and sodium (23) intakes. Details on the laboratories and methods used are provided in Web Appendix 1 (available at https://academic.oup.com/aje).

Urinary nitrogen level (in grams) was divided by 0.81 to convert the data to dietary nitrogen values (21) and then multiplied by 6.25 to convert the data to dietary protein values. Urinary potassium level was divided by 0.8 to convert the data to dietary potassium values (24), and urinary sodium level was divided by 0.86 to convert the data to dietary sodium values (25).

The timing of measurements varied across studies (Web Figure 1). The FFQ was administered at entry in 4 studies but 1–14 months after entry in the AMPM study. In all studies, doubly labeled water was measured at 1–14 days, and 24-hour urine samples were collected during the same period. The 24HRs were administered on the following days: OPEN, days 1 and 61; Energetics, days 1, 3, 5, 8, 11, 14, 30, and 60; AMPM, days 1, 5–6, and 10–11; NBS, around 180 days (twice in a 20% subsample); NPAAS, days 14–104 (3 administrations).

These measurements were supplemented by substudies, of varying size, carried out to examine the reliability of self-reports and biomarkers. The duration of time between main and substudy administrations ranged from 2 weeks in OPEN to approximately 6 months in Energetics, NBS, and NPAAS and 10–23 months in AMPM. In OPEN, only the doubly labeled water measurement was repeated; in other studies, all biomarker measurements and self-reports were repeated. For example, in NBS and NPAAS, the entire study protocol was repeated in a 20% subsample. Our analysis included repeat biomarker and 24HR assessments.

Statistical methods

We report on 7 dietary components: energy intake and intakes of protein, potassium, and sodium and their densities (ratios to energy intake).

We predicted targeted true usual intake (defined here as the 12-month average) using different self-report instruments. A measure of the goodness of a prediction is its correlation with truth. The correlation measures how well the prediction orders individuals according to their true intake and relates to loss of statistical power (26). Low values (e.g., <0.4) are undesirable, although there is no sharp cutoff.

The time-varying intake model accounts for serial correlation in individuals' intakes and proximity of self-report assessments to biomarker assessments (9). We describe the model briefly below. For technical details, see Web Appendix 2 and Freedman et al. (9).

Each sex and dietary component was modeled separately. Data on dietary variables were logarithmically transformed. The model comprised 4 meta-analysis submodels with study-specific parameters. The first 3 submodels specified linear regression relationships between the biomarker, 24HR, and FFQ, respectively, and true intake, the explanatory variable. The fourth submodel specified how true intake varied over time. The time axis was divided into 90-day subperiods.

Biomarker submodel. In the biomarker submodel, biomarkers were assumed to measure true intake on the previous day (or for energy during the 10- to 14-day assessment) without bias (intercept = 0, slope = 1) but with independent random error. The error variance was estimated through repeat assessments performed within the same subperiod and was assumed to be equal across studies, because of insufficient replications to provide study-specific estimates.

Table 1. Correlation Coefficients for the Correlation of Intakes of Selected Dietary Components (Predicted[a] From Various Combinations of Self-Report Instruments[b]) With Truth, Combined Over 5 Validation Studies[c]

Dietary Component, Sex, and Use of FFQ (Y/N)[d]	No. of 24HRs Administered				
	0	1	2	4	8
Energy intake					
Male					
Y	0.15	0.28	0.32	0.36	0.38
N		0.28	0.33	0.36	0.39
Female					
Y	0.22	0.24	0.25	0.26	0.28
N		0.18	0.22	0.24	0.27
Protein intake					
Male					
Y	0.30	0.42	0.48	0.53	0.57
N		0.37	0.45	0.52	0.56
Female					
Y	0.35	0.47	0.52	0.57	0.62
N		0.40	0.49	0.56	0.61
Potassium intake					
Male					
Y	0.48	0.54	0.58	0.60	0.63
N		0.44	0.52	0.57	0.61
Female					
Y	0.39	0.53	0.58	0.63	0.66
N		0.47	0.56	0.62	0.66
Sodium intake					
Male					
Y	0.21	0.36	0.43	0.48	0.53
N		0.34	0.42	0.48	0.53
Female					
Y	0.19	0.22	0.24	0.27	0.30
N		0.17	0.22	0.26	0.29
Protein density					
Male					
Y	0.40	0.47	0.51	0.56	0.60
N		0.35	0.44	0.52	0.59
Female					
Y	0.45	0.46	0.47	0.47	0.48
N		0.27	0.34	0.39	0.44
Potassium density					
Male					
Y	0.48	0.53	0.57	0.61	0.64
N		0.44	0.53	0.60	0.65
Female					
Y	0.50	0.54	0.56	0.57	0.58
N		0.39	0.46	0.52	0.56
Sodium density					
Male					
Y	0.36	0.46	0.52	0.57	0.63
N		0.38	0.47	0.56	0.62
Female					
Y	0.34	0.36	0.37	0.39	0.41
N		0.20	0.26	0.34	0.40

Abbreviations: FFQ, food frequency questionnaire; 24HR, 24-hour recall; N, no; Y, yes.

[a] The predictions were formed under the assumption that 24HRs are unbiased measures of intake.

[b] An FFQ, 24HRs (1, 2, 4, or 8), or a combination of these.

[c] The Observing Protein and Energy Nutrition Study (1999–2000), the Energetics Study (2006–2009), the Automated Multiple-Pass Method validation study (2002–2004), the Nutrition Biomarker Study (2004–2005), and the Nutrition and Physical Activity Assessment Study (2007–2009).

[d] Whether or not an FFQ report was used in predicting usual intake.

24HR submodel. In the 24HR submodel, 24HR-reported intake was assumed to measure true intake on the same day with systematic intake-related bias. Person-specific bias and within-person random-error terms were included, as in Kipnis et al. (27). Within-person random-error terms were assumed to be mutually independent and independent of all other terms in the model.

FFQ submodel. In the FFQ submodel, FFQ-reported intake was assumed to measure true average intake over the past year with systematic intake-related bias. With no FFQ replications, person-specific bias and within-person random-error terms could not be separately estimated and were combined as a total error term. This total error was assumed to be correlated with the 24HR person-specific bias.

Time-varying true intake submodel. In the time-varying true intake submodel, we assumed that 1) an individual's true intake varied over time, but the group average and variance, on a single day and in each subperiod, remained constant within-study; 2) the ratio of the single-day intake variance to the 90-day average intake variance (the intake-variance ratio) was common across studies; and 3) the correlation structure between usual intakes in different subperiods was common across studies and was autoregressive (AR) of order 1 (AR(1)) or compound symmetry, decided according to the best model fit, using Akaike's Information Criterion.

Model parameters were estimated by means of maximum likelihood, assuming log biomarker, 24HR, and FFQ values to be normally distributed, using the CALIS procedure in SAS (SAS Institute, Inc., Cary, North Carolina) (28).

Instruments were combined through a prediction (or calibration) equation for dietary intake, using 24HRs as the reference method. Via estimated model parameters, we estimated correlations of these predictions with truth, separately for each study,

sex, and dietary component, for 24HRs, the FFQ, or a combination. We present results for a single FFQ and 1, 2, 4, or 8 24HRs with or without an FFQ. The estimates depend on the timing of self-reports. The timing assumed is shown in Web Table 1, and the calculation methods are given in Web Appendix 2. Study-specific estimates are presented in the Web material (specified below), and across-study summaries (weighted averages) are presented in the main paper.

RESULTS

Details on the models fitted and parameter estimates are provided in Web Tables 2–12.

Table 1 shows, for each dietary component, the estimated correlations with truth, averaged across studies, for a single FFQ, different numbers of 24HR administrations, and their combination. Correlations increased with the number of 24HR administrations, but with diminishing benefits. Similarly, adding an FFQ to one or more 24HRs improved correlations, but benefits diminished with increasing number of 24HRs. Benefits of combining 24HRs with an FFQ were seen for both absolute intakes and densities, but they appeared smaller for women than for men. Estimated correlations remained low (<0.40) for energy intake among men and women and for sodium intake and sodium density among women. For other dietary components, correlations above 0.5 were attained for some combinations, but for protein density in women the maximum was 0.48. Study-specific correlations are presented in Web Tables 13–20.

We considered the "least" combination (the combination with the smallest number of self-report administrations) yielding a correlation at least 90% of that achieved by using 1 FFQ plus 8 24HRs (Table 2). One FFQ plus 3 or 4 24HRs provided the majority of the benefit derived from combining instruments. See Web Figures 2–15 for graphs of the relative correlations achieved by using different combinations of instruments.

Using multiple 24HRs with an FFQ, the predicted dietary intake used in the analysis also provides estimated risk parameters that have less attenuation due to measurement error. As an example, using 1 FFQ plus 4 administrations of a 24HR to predict dietary intake, the attenuation coefficients are low for energy, are considerably higher for absolute intakes of protein and potassium than the 0.3–0.4 values typically seen when no adjustment for measurement error is made, and are quite close to 1.0 (no attenuation) for nutrient densities (Web Table 21).

Table 2. "Least" Combination[a] of an FFQ and 24HRs Achieving a Correlation of at Least 90% of the Correlation for 1 FFQ + 8 24HRs in Assessment of Selected Dietary Intakes, by Sex, in 5 Validation Studies[b]

Dietary Component	No. of FFQs + No. of 24HRs	
	Men	Women
Energy intake	0 + 3	1 + 2
Protein intake	1 + 3	1 + 3
Potassium intake	1 + 2	1 + 3
Sodium intake	0 + 4	1 + 4
Protein density	1 + 4	1 + 0
Potassium density	1 + 3	1 + 1
Sodium density	1 + 4	1 + 2

Abbreviations: FFQ, food frequency questionnaire; 24HR, 24-hour recall.

[a] The combination with the smallest number of self-report administrations yielding a correlation at least 90% of that achieved by using 1 FFQ plus 8 24HRs.

[b] The Observing Protein and Energy Nutrition Study (1999–2000), the Energetics Study (2006–2009), the Automated Multiple-Pass Method validation study (2002–2004), the Nutrition Biomarker Study (2004–2005), and the Nutrition and Physical Activity Assessment Study (2007–2009).

DISCUSSION

Adding repeat administrations of a 24HR to an FFQ increased the accuracy of prediction of dietary intake, but with diminishing returns as the number of 24HR administrations increased. One FFQ plus 4 24HR administrations gave at least 90% of the maximum correlation. Using this combination, correlations between predicted and true intakes were 0.39–0.61 for nutrient densities compared with 0.35–0.51 for a single FFQ. Absolute increases in correlation ranged from 0.02 to 0.26, depending on nutrient and sex, with an average increase of 0.14. These results strengthen those of Carroll et al. (5), being based on unbiased recovery biomarker evaluation.

Improvements in correlation from adding a single FFQ to 4 administrations of a 24HR were mostly marginal. However, the dietary components we studied are not episodically consumed. Carroll et al. (5) also reported only small increases in correlations for such components.

The correlations obtained by combining 24HRs with an FFQ leave room for considerable improvement. Clearly, continued research into ways of assessing dietary intake is needed, so as to increase correlations to more than 0.7.

Neuhouser et al. (15) showed that for energy and absolute protein intake, correlations can be considerably increased by including in the prediction equation body mass index (weight (kg)/height $(m)^2$) and other personal characteristics that are sources of systematic reporting error. More extensive data analysis confirms this and extends the result to absolute sodium intake and potassium density, with modest gains for potassium intake, protein density, and sodium density (8, 11). Such prediction equations hold promise for further gains in the precision of intake measurement.

However, such prediction is effective only when it is derived from biomarker references. Here, we have focused on studies collecting data from an FFQ and 24HRs, where prediction equations are based on 24HRs as reference instruments. Such equations do not capture the importance of personal characteristics such as body mass index (29). Consequently, prediction equations based on 24HRs as the reference method do not greatly improve correlations with truth (Web Tables 22 and 23). Biomarker-based prediction equations are preferable, but they require cohort studies to include validation substudies with such biomarkers and are limited to the few nutrients with suitable biomarkers.

A central assumption behind our modeling is that recovery biomarkers are unbiased for individual intake and that their errors are random. We discuss this assumption in Web Appendix 2.

The results presented here are for the limited set of dietary components that have recovery biomarkers. Carroll et al. (5) have shown that combining 24HRs and an FFQ can lead to improvements across a wider range of dietary components. Together, this suggests that while the gains we have demonstrated will not apply to all dietary components, they will apply to many nutrients and foods.

ACKNOWLEDGMENTS

Author affiliations: Biostatistics Unit, Gertner Institute for Epidemiology, Sheba Medical Center, Tel Hashomer, Israel (Laurence S. Freedman); Information Management Services, Inc., Rockville, Maryland (Laurence S. Freedman); Biometry Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, Maryland (Douglas Midthune, Victor Kipnis); Division of General Internal Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California (Lenore Arab); Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington (Ross L. Prentice, Marian L. Neuhouser, Lesley F. Tinker); Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland (Amy F. Subar); Departments of Epidemiology and Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Walter Willett); Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts (Walter Willett); and Division of Nutrition, Harvard Medical School, Boston, Massachusetts (Walter Willett).

This work was supported by the National Heart, Lung, and Blood Institute (contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C for the Women's Health Initiative) and the National Cancer Institute (grant R01 CA119171 for the Nutrition and Physical Activity Assessment Study).

We thank the following people for their valuable contributions to this work: Dr. Alfonso Ang (University of California, Los Angeles, Los Angeles, California), John Commins (Information Management Services, Inc., Rockville, Maryland), Dr. Sheila Bingham (deceased; Dunn Nutrition Unit, Medical Research Council, Cambridge, United Kingdom), Dr. Kevin Dodd (National Cancer Institute, Bethesda, Maryland), Dr. Alanna Moshfegh (US Department of Agriculture, Washington, DC), Dr. Nancy Potischman (National Institutes of Health, Bethesda, Maryland), Dr. Dale Schoeller (University of Wisconsin, Madison, Wisconsin), and Dr. Richard Troiano (National Institutes of Health). We also thank the Women's Health Initiative investigators and staff for their dedication. A full listing of Women's Health Initiative investigators can be found at https://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf.

L.A. has an intellectual property interest in DietDay (Centrax Corporation, Chicago, Illinois), the Web-based 24-hour recall instrument used in the Energetics Study. None of the other authors have any conflicts of interest to declare.

REFERENCES

1. Subar AF, Freedman LS, Tooze JA, et al. Addressing current criticism regarding the value of self-report dietary data. *J Nutr*. 2015;145(12):2639–2645.
2. Boushey CJ, Spoden M, Zhu FM, et al. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. *Proc Nutr Soc*. 2017;76(3):283–294.
3. Stumbo PJ. New technology in dietary assessment: a review of digital methods in improving food record accuracy. *Proc Nutr Soc*. 2013;72(1):70–76.
4. Freedman LS, Midthune D, Carroll RJ, et al. Using regression calibration equations that combine self-reported intake and biomarker measures to obtain unbiased estimates and more powerful tests of dietary associations. *Am J Epidemiol*. 2011;174(11):1238–1245.
5. Carroll RJ, Midthune D, Subar AF, et al. Taking advantage of the strengths of two different dietary assessment instruments to improve intake estimates for nutritional epidemiology. *Am J Epidemiol*. 2012;175(4):340–347.
6. Subar AF, Thompson FE, Kipnis V, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. *Am J Epidemiol*. 2001;154(12):1089–1099.
7. Kaaks RJ. Biochemical markers as additional measurements in studies of the accuracy of dietary questionnaire measurements: conceptual issues. *Am J Clin Nutr*. 1997;65(4 suppl):1232S–1239S.
8. Freedman LS, Commins JM, Moler JE, et al. Pooled results from five validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. *Am J Epidemiol*. 2014;180(2):172–188.
9. Freedman LS, Midthune D, Dodd KW, et al. A statistical model for measurement error that incorporates variation over time in the target measure, with application to nutritional epidemiology. *Stat Med*. 2015;34(27):3590–3605.
10. Freedman LS, Midthune D, Carroll RJ, et al. Application of a new statistical model for measurement error to the evaluation of dietary self-report instruments. *Epidemiology*. 2015;26(6):925–933.
11. Freedman LS, Commins JM, Moler JE, et al. Pooled results from five validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. *Am J Epidemiol*. 2015;181(7):473–487.
12. Subar AF, Kipnis V, Troiano RP, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN Study. *Am J Epidemiol*. 2003;158(1):1–13.
13. Moshfegh AJ, Rhodes DG, Baer DJ, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *Am J Clin Nutr*. 2008;88(2):324–332.
14. Arab L, Wesseling-Perry K, Jardack P, et al. Eight self-administered 24-hour dietary recalls using the Internet are feasible in African Americans and Caucasians: The Energetics Study. *J Am Diet Assoc*. 2010;110(6):857–864.

15. Neuhouser ML, Tinker L, Shaw PA, et al. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol*. 2008;167 (10): 1247–1259.
16. Prentice RL, Mossavar-Rahmani Y, Huang Y, et al. Evaluation and comparison of food records, recalls and frequencies for energy and protein assessment by using recovery biomarkers. *Am J Epidemiol*. 2011;174(5):591–603.
17. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol*. 1985;122(1): 51–65.
18. Patterson RE, Kristal AR, Tinker LF, et al. Measurement characteristics of the Women's Health Initiative food frequency questionnaire. *Ann Epidemiol*. 1999;9(3): 178–187.
19. National Heart, Lung, and Blood Institute. Women's Health Initiative. Food Questionnaire. (Form 60). https://www.whi.org/researchers/studydoc/_layouts/15/WopiFrame.aspx?sourcedoc=/researchers/studydoc/WHI%20Forms/F060%20v1.6.pdf&action=default. Published 1993. Last updated May 2003. Accessed June 4, 2018.
20. Schoeller DA, Hnilicka JM. Reliability of the doubly labeled water method for the measurement of total daily energy expenditure in free-living subjects. *J Nutr*. 1996;126(1): 348S–354S.
21. Bingham SA, Cummings JH. Urine nitrogen as an independent validatory measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet. *Am J Clin Nutr*. 1985;42(6):1276–1289.
22. Mickelsen O, Makdani D, Gill JL, et al. Sodium and potassium intakes and excretions of normal men consuming sodium chloride or a 1:1 mixture of sodium and potassium chlorides. *Am J Clin Nutr*. 1977;30(12):2033–2040.
23. Luft FC, Fineberg NS, Sloan RS. Estimating dietary sodium intake in individuals receiving a randomly fluctuating intake. *Hypertension*. 1982;4(6):805–808.
24. Freedman LS, Midthune D, Carroll RJ, et al. Adjustments to improve the estimation of usual dietary intake distributions in the population. *J Nutr*. 2004;134(7):1836–1843.
25. Holbrook JT, Patterson KY, Bodner JE et al. Sodium and potassium intake and balance in adults consuming self-selected diets. *Am J Clin Nutr*. 1984;40(4):786–793.
26. Kaaks R, Riboli E, van Staveren W. Calibration of dietary intake measurements in prospective cohort studies. *Am J Epidemiol*. 1995;142(5):548–556.
27. Kipnis V, Subar AF, Midthune D, et al. The structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol*. 2003;158(1):14–21.
28. SAS Institute Inc. *SAS Software, Version 9.2*. Cary, NC: SAS Institute Inc.; 2008.
29. Freedman LS, Commins JM, Willett W, et al. Evaluation of the 24-hour recall as a reference instrument for calibrating other self-report instruments in nutritional cohort studies: evidence from the Validation Studies Pooling Project. *Am J Epidemiol*. 2017;186 (1):73–82.