Application of a New Statistical Model for Measurement Error to the Evaluation of Dietary Self-report Instruments

HHS Public Access

Author manuscript

Epidemiology. Author manuscript; available in PMC 2016 June 08.

Published in final edited form as:

Epidemiology. 2015 November ; 26(6): 925–933. doi:10.1097/EDE.0000000000000377.

Application of a New Statistical Model for Measurement Error to the Evaluation of Dietary Self-report Instruments

Laurence S. Freedman[a], **Douglas Midthune**[b], **Raymond J. Carroll**[c], **John M. Commins**[a], **Lenore Arab**[d], **David J. Baer**[e], **James E. Moler**[a], **Alanna J. Moshfegh**[e], **Marian L. Neuhouser**[f], **Ross L. Prentice**[f], **Donna Rhodes**[e], **Donna Spiegelman**[g], **Amy F. Subar**[b], **Lesley F. Tinker**[f], **Walter Willett**[g], and **Victor Kipnis**[b]

[a]Information Management Services, Inc., Rockville, MD

[b]National Cancer Institute, Bethesda, MD

[c]University Texas A&M, College Station, TX

[d]University of California Los Angeles, Los Angeles, CA

[e]United States Department of Agriculture, Beltsville, MD

[f]Fred Hutchinson Cancer Research Center, Seattle, WA

[g]Harvard School of Public Health, Boston, MA

Abstract

Most statistical methods that adjust analyses for dietary measurement error treat an individual's usual intake as a fixed quantity. However, usual intake, if defined as average intake over a few months, varies over time. We describe a model that accounts for such variation and for the proximity of biomarker measurements to self-reports within the framework of a meta-analysis, and apply it to the analysis of data on energy, protein, potassium, and sodium from a set of five large validation studies of dietary self-report instruments using recovery biomarkers as reference instruments. We show that this time-varying usual intake model fits the data better than the fixed usual intake assumption. Using this model, we estimated attenuation factors and correlations with true longer-term usual intake for single and multiple 24-hour dietary recalls (24HRs) and food frequency questionnaires (FFQs) and compared them with those obtained under the "fixed" method. Compared with the fixed method, the estimates using the time-varying model showed slightly larger values of the attenuation factor and correlation coefficient for FFQs and smaller values for 24HRs. In some cases, the difference between the fixed method estimate and the new estimate for multiple 24HRs was substantial. With the new method, while four 24HRs had higher estimated correlations with truth than a single FFQ for absolute intakes of protein, potassium, and sodium, for densities the correlations were approximately equal. Accounting for the time element in dietary validation is potentially important, and points toward the need for longer-term validation studies.

Correspondence: Laurence S. Freedman, Gertner Institute for Epidemiology, Sheba Medical Center, Tel Hashomer 52621, Israel. ; Email: lsf@actcom.co.il

An extensive literature exists on statistical methods for dealing with dietary measurement error. Most methods specify a model linking an individual's self-reported intake to his/her true usual intake, which is treated as a fixed quantity.[1]

However, usual or average intake in dietary research is often not defined clearly. A precise definition would require specifying the period over which the average is taken, but often such specification is absent. This can lead to vagueness of definition in measures of accuracy of self-report instruments. For example, consider the correlation with true usual intake of reported intakes from multiple 24-hour recalls (24HRs) taken over 2 weeks. This correlation may vary according to whether usual intake is defined as the average over the month, 3 months, year, or several years that are proximal to the time of the recalls. Clearly, the longer the usual intake period, the lower the expected correlation is between the report and usual intake. This is because dietary intakes on any 2 days tend to be closer, the closer are the 2 days in time[2] (cyclical variations between weekdays and weekends and between seasons excepted).

Three notable exceptions to the fixed usual intake approach are described by Rosner et al.,[3] Keogh et al.,[4] and Prentice and Huang.[5] We discuss these approaches in the eAppendix, Supplemental Digital Content 1, at the end of the section entitled "Statistical analysis, model and estimation of parameters" (http://links.lww.com/EDE/A967).

In this article, we describe a model where short-term usual (i.e., average) intake varies from one short-term period (we use 3 months) to the next. The targeted longer-term usual intake is then an average of several short-term usual intakes (we choose four, giving a targeted usual intake period of 1 year). The modeling requires (1) assuming no systematic trend in average intake over the targeted period and (2) estimating the correlation between intakes in any two separate short-term periods. However, in any single study, there are often only two repeats chosen to be approximately equally spaced, thus limiting the correlations that can be estimated. To overcome this, we analyze several different studies, each of which uses a different period between repeat biomarker evaluations; overall, we are thus able to cover the targeted 1-year period. We therefore describe our model within a meta-analysis framework, so as to apply it to data that come from the Validation Studies Pooling Project (VSPP),[6,7] comprising five large dietary validation studies that used recovery biomarkers.[7] In the "Methods" section, we describe the VSPP and the statistical model and methods. In the "Results" section, we describe the results of applying the method to VSPP data and compare them with results obtained assuming a fixed usual intake. In the "Discussion" section, we discuss the implications of our results.

METHODS

The Validation Studies Pooling Project

Dietary intake recovery biomarkers[8] that provide accurate assessments of short-term intakes provide the most acceptable method of evaluating dietary self-report instruments.[9] However, these biomarkers are expensive or inconvenient, and exist for only a limited set of dietary components (energy, protein, potassium, and sodium). In 2009, investigators of five larger

(>200 participants) validation studies using such biomarkers agreed to pool their data for common analysis. The resulting VSPP aims to clarify the nature and magnitude of reporting errors in food frequency questionnaires (FFQs) and 24HRs.[6,7]

The five VSPP studies included diverse populations within the US.[6] The Observing Protein and Energy study (OPEN)[10] and the Automated Multiple Pass Method (AMPM) validation study[11] both included adult volunteers, ages 40–69 years, residing in Maryland. The Energetics study included younger white and African-American adults residing in California.[12] The Nutrition Biomarker Study (NBS) included women participants in the Dietary Modification Trial,[13] and the Nutrition and Physical Activity Assessment Study (NPAAS) women participants in the Observational Cohort of the Women's Health Initiative.[14] These participants were mostly over 60 years, residing throughout the US. Further details are in references 9–13.

Each study included administering a FFQ to each participant. Although repeat administrations were sometimes performed, this analysis includes only the first administration. The FFQs queried intake over the past year in OPEN, Energetics and AMPM, and the past 3 months in NBS and NPAAS. Three versions of FFQ were used[15_18] (see eAppendix, Supplemental Digital Content 1, Section entitled "Further details of the design of the VSPP studies"; http://links.lww.com/EDE/A967).

Each study included two or more 24HR assessments, administered to all participants in four studies, and to a 20% subset in NBS. Different versions of 24HR were used[10_19] (see eAppendix, Supplemental Digital Content 1, Section entitled "Further details of the design of the VSPP studies"; http://links.lww.com/EDE/A967).

Each study included the recovery biomarkers: doubly-labeled water for energy intake,[20] 24-hour urinary nitrogen for protein,[21] 24-hour urinary potassium for potassium,[22] and 24-hour urinary sodium for sodium.[23]

Doubly-labeled water measures energy expenditure over a 10- to 14-day period and, assuming individuals are in energy balance, is used to measure average daily energy intake.[20]

The 24-hour urinary markers assess intake over a 24-hour period.[22,23] For details of the methods and laboratories, see eAppendix, Supplemental Digital Content 1, Section entitled "Further details of the design of the VSPP studies" (http://links.lww.com/EDE/A967). Three studies included repeat determinations in the main protocol, approximately 5 days apart; NBS and NPAAS included repeat determinations in a 20% sub-study (see below), about 6 months later. The timing of the different measurements varied across studies (see eAppendix, Supplemental Digital Content 1, Section entitled "Further details of the design of the VSPP studies"; http://links.lww.com/EDE/A967).

Urinary nitrogen in grams was divided by 0.81 to convert to dietary nitrogen,[21] and multiplied by 6.25 to convert dietary nitrogen to dietary protein. Urinary potassium was divided by 0.8 to convert to dietary potassium,[24] and urinary sodium by 0.86 to convert to dietary sodium.[25]

Each study included a sub-study, of varying size, to examine reliability of self-reports and biomarkers. The time between main and sub-study administrations ranged from 2 weeks in OPEN, to approximately 6 months in Energetics, NBS, and NPAAS, and to 10–23 months in AMPM. The extent of the substudy data collection also varied. In OPEN, only doubly-labeled water was repeated, while other studies repeated biomarkers and self-reports. For example, NBS and NPAAS repeated the entire study protocol in a 20% subsample. Our analysis included repeat biomarker and 24HR assessments.

Statistical Methods

We report on seven dietary components: energy, protein, potassium, sodium, protein density, potassium density, and sodium density. Protein density is defined as the percent of total energy from protein; potassium and sodium densities are defined as the ratio of nutrient intake (mg) to energy intake (1,000 kcal).

When relating longer-term average intake of some dietary component to a health outcome, we should be interested in how well our self-report instruments capture longer-term average intake. Here, we choose to target a 12-month average, although our method is easily adapted to other periods within the compass of the data collected. We call this 12-month average the *targeted true usual intake*.

Two important measures of a self-report instrument are the attenuation factor and the correlation with targeted true usual intake. The attenuation factor (usually between 0 and 1) is the multiplicative bias or shrinkage factor in the estimated regression coefficient when a health outcome is regressed on continuous self-reported rather than true usual intake.

The correlation coefficient between reported and true intake is used to measure loss of statistical power to detect diet–health associations when using reported instead of true intake.[26] In simple models, it can also serve to de-attenuate relative risks between two intake categories.[27] Low values of attenuation factor and correlation, e.g., less than 0.4, are undesirable, although there is no sharp cut-off. An attenuation factor of 0.4 leads to a true relative risk of 2.0 being attenuated to $2^{0.4} = 1.32$.

We now describe estimating attenuation factors and correlations, using the time-varying usual intake model.

Statistical Modeling

In four of the five studies, participants completed the first FFQ at the beginning of the study, and in the remaining study toward its end. We set the completion of this FFQ as the common time point. Relative to this FFQ, other instruments were completed from 450 days beforehand to 450 days afterwards. We divided this overall period into 10 subperiods of 90 days each, with the FFQ being completed at the beginning of the 6th period (Figure 1).

For modeling, each type of observation, 24HR, FFQ, and biomarker was considered an error prone measurement of true intake. A statistical model describes how each of them relates to true intake.

We used the same model for each sex and each dietary component, and the modeling of each sex and component was performed separately. All dietary variables were logarithmically transformed, including the unobserved true intake. We excluded urinary marker values from analysis only if participants declared missing two or more voids during 24-hour collection.[28] For more details on this, on exclusion of outliers, and for precise details of the statistical model see eAppendix, Supplemental Digital Content 1, Section entitled "Statistical analysis, model and estimation of parameters" (http://links.lww.com/EDE/A967).

The model has four parts; each is a meta-analysis model that specifies study-specific parameters. The first three parts specify a linear regression relationship between the biomarker, 24HR and FFQ, respectively, and true intake; the instrument is the dependent variable and true intake the explanatory variable. The fourth part specifies how true intake varies over time.

Biomarker Model—Biomarkers were assumed to measure true intake on the day of the assessment (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 90-day subperiod.

24HR Model—24HR-reported intake was assumed to measure true intake on the day before assessment with constant systematic and also intake-related bias. A person-specific bias and a within-person random error term were included, as in Kipnis et al.[29] For developing calibration equations for predicting true usual intake, extra covariates representing personal characteristics such as log body mass index (BMI) were introduced as linear terms. The within-person random error terms were assumed mutually independent, and independent of true intake, all person-specific biases, and within-person random errors in other parts of the model.

FFQ Model—FFQ-reported intake was assumed to measure true average intake over the past year with constant systematic and also intake-related bias. A person-specific bias and a within-person random error term were included. The person-specific bias terms of FFQ and 24HR could be correlated. As with the 24HR, covariates could be included as extra linear terms, depending on the purpose of the analysis. Theoretically, for the regression calibration approach to measurement error adjustment, covariates that are included in the disease model should also be included in such a prediction.[1] However, it has been customary to omit extra covariates when estimating attenuation factors and correlations. We examined both.

Time-varying True Intake Model—A stochastic structure for true intake was required. We assumed that:

1. a person's true intake varied over time, but the group average and variance, on a single day and in each 90-day subperiod, remained constant for a given study;
2. the ratio of the single-day variance to the 90-day usual intake variance, which we call the intake–variance ratio, was common across studies;

3. the correlation structure between usual intakes in different subperiods was common across studies, and was autoregressive of order 1, compound symmetry, or degenerate (all 1's). The autoregressive of order 1 structure has correlation between two subperiods j and $k = \rho^{|j-k|}$, where $-1 \leq \rho \leq 1$; with compound symmetry this correlation equals ρ. The structure was decided according to the best model fit, using Akaike's information criterion. The degenerate option, together with assuming that the intake-variance ratio equals 1, corresponds to the fixed usual intake (no variation over time) model. We used this to compare estimates obtained under the fixed versus the time-varying model.

We assumed study-specific model parameters, with the following exception. The biomarker error variance was assumed equal for each study, since there were insufficient replications of some biomarker measurements within subperiods to provide study-specific estimates. Model parameters were estimated using maximum likelihood, assuming log biomarker, 24HR, and FFQ values to be normally distributed (conditional on covariates, if included). However, when first and second moments are correctly specified, parameters are consistently estimated without the normality assumption. Using the nonparametric bootstrap for estimating standard errors also obviates the need to assume normality. Estimation was performed using a custom-built SAS[30] program (eAppendix, Supplemental Digital Content 1, Section entitled "Examples of SAS code"; http://links.lww.com/EDE/A967).

We estimated attenuation factors and correlations with longer-term true usual intake for each study, sex, instrument, and dietary component via the estimated model parameters. See eAppendix, Supplemental Digital Content 1, Section entitled "Statistical analysis, model and estimation of parameters" (http://links.lww.com/EDE/A967) for the equations. Weighted averages of attenuation factors and correlations across the five studies were calculated, with weights inversely proportional to the nonparametric bootstrap variance of the estimates.

Model fit was investigated using Akaike's information criterion, and comparing empirical correlations between repeat biomarkers with those predicted from the model.

Calibration Equations

Calibration equations for predicting the targeted true usual intake using self-report and personal characteristics were calculated from parameter estimates of the model. For more details, see eAppendix, Supplemental Digital Content 1, Section entitled "Calibration equations" (http://links.lww.com/EDE/A967). We illustrate the method using FFQ-reported intakes, and the characteristics: age group (<40 years, 40–49, 50–59, 60–69, 70–79, ≥80), log BMI, race (African-American vs. other), and education (high school, college, postgraduate). Comments on choosing covariates in the calibration model are given in the "Discussion."

RESULTS

Model Fit

For each sex and dietary component, the correlation structure with the best model fit (lowest Akaike information criterion) is shown in Table 1. Also presented, are estimates of the

correlation parameter ρ and the intake–variance ratio (see “Methods”). In most cases, the autoregressive structure provided the best fit, which, comparing Akaike information criterion values, was far superior to that provided by the fixed intake model.

Examples of comparing empirical correlations between repeat biomarker measurements with their model-predicted values are shown for energy and protein density in Figures 2 and 3. Many of the empirical correlations were estimated with less than 20 pairs of observations, so error bars were wide. The model-based estimates fell within the error bars in all cases except one (protein density for women, measurements four subperiods apart), meaning that the model appeared consistent with the data.

Attenuation Factors and Correlations with Targeted Usual Intake

Tables 2 and 3 present attenuation factors and correlations with targeted usual intake, estimated from the fixed and time-varying intake models. Across-study average estimates are provided for FFQs, single 24HRs, and the average of four 24HRs.

The same pattern was observed for all dietary components and both sexes. Estimated attenuation factors and correlations for FFQs were similar or slightly larger under the time varying than under the fixed model. However, for 24HRs, the estimates were smaller under the time varying than under the fixed model. The difference between estimates was larger for the average of four 24HRs, and was substantial for women reporting protein density and sodium density. This pattern was quite consistent across the five studies.

As a result, whereas under the fixed model, four 24HRs had larger estimated attenuation factors and correlations than a single FFQ for all dietary components, under the time-varying model this was true for absolute intakes of protein, potassium, and sodium, but not their densities. For these densities, under the time-varying intake model, estimated attenuation factors and correlations for four 24HRs appeared similar to those for a single FFQ.

Calibration Equations

The coefficients of personal characteristics variables in FFQ-based calibration equations estimated from the time-varying model (see eAppendix, Supplemental Digital Content 1, eTables A1–A6; http://links.lww.com/EDE/A967) were similar to those for the fixed model (see references 6, 7). However, the coefficient of the FFQ-report estimated from the time-varying model displayed the same pattern as did the attenuation factors in relation to estimates based on the fixed model. For example, with a single 24HR, the time-varying model coefficient estimates were on average 21% lower than those from the fixed model. Similarly the estimated R^2 values were on average 14% lower under the time-varying model than under the fixed model. For FFQs, there was little difference between the calibration equations estimated under the two models.

With either model, the covariates that were important for predicting intakes were: for energy and protein: age, BMI, and race; for protein density: age and race; for potassium: age, race, and education; for potassium density: age, BMI, race, and education; for sodium: age and BMI; and for sodium density: BMI and race. Multiple correlations were considerably higher

for the full calibration model compared with the instrument-only model for energy, protein and sodium, but the difference was smaller for potassium and the nutrient densities.

DISCUSSION

We have described a time-varying usual intake model for analyzing dietary validation data that incorporates different types of instrument and accounts for the relative timing of biomarker measurements to self-report instruments and the time between repeated measurements. Applying the method to VSPP data showed that this model provided a better fit to the data than assuming a fixed intake, and led to reductions in estimates of attenuation factor and correlation with targeted true usual intake for 24HRs, sometimes substantially, compared with the fixed intake method. We think this difference in estimates was related to a design feature of some of the validation studies (see below).

Preis et al.[31] reported investigating whether the timing of repeat biomarkers influences estimates of attenuation factor and correlation. They used data from OPEN and AMPM, and a fixed intake model, and applied estimates of within-person biomarker variance found in AMPM to OPEN. They reported that estimated correlations decreased when based on biomarker repeats close in time and increased when based on repeats spaced further apart, although this was disputed by Dodd et al.[32] Our results in Tables 2 and 3 showed a different pattern, where estimated correlations for FFQs were relatively unaffected by using a fixed model, whereas those for 24HRs were overestimated.

Apparently, this differential result was mostly due to a design feature of some of our studies, performing biomarker measurements a day before 24HRs, thus measuring the same day's intake. If biomarker and 24HR assessments had not been so timed, it seems likely that 24HR correlations estimated by the fixed method would have had less bias. However, both proximity of self-report to biomarker determination and time between repeat biomarker measurements influence the estimates of attenuation factor and correlation using the fixed method, and both should be accounted for using a time-varying intake model.

The pattern of differences between the time-varying and fixed model estimates was the same across all of the nutrients examined. It therefore seems likely that, were recovery biomarkers available for other nutrients, one would see the same pattern.

As mentioned in the "Introduction," there is previous study[3–5] that did not assume a fixed usual intake value. Some details of each are provided in the eAppendix, Supplemental Digital Content 1, end of Section entitled "Statistical analysis, model and estimation of parameters" (http://links.lww.com/EDE/A967). Our proposed model is related to these approaches, but, unlike them, specifically models the correlation of true intake across time and estimates attenuation factors and correlations for a longer-term average intake.

A central assumption behind our modeling is that recovery biomarkers are unbiased for individual usual intake, and that errors in their measurements are random. There is a literature supporting this claim (see references 33, 34 for reviews), although some discussion continues regarding urinary potassium and sodium. Freisling et al.[35] reviewed the literature and reported conversion factors for potassium that varied from 0.76 to 0.89. Although there

is not complete agreement, the appropriate conversion factor, in the context of this article, this choice does not impact our results. As we use the log biomarker level, the conversion factor translates to a simple additive constant that does not affect estimates of attenuation factors or correlations.

More recently, Turban et al.[36] have reported that potassium excretion was lower among black than white participants in the Dietary Approaches to Stop Hypertension (DASH) trial who were fed the high sodium diet. This finding challenges the conventional assumption that the fraction of potassium intake excreted in urine is independent of personal characteristics, and further investigations (for example, feeding studies) are warranted. For a discussion of the urinary sodium biomarker, see the "Discussion" section of Freedman et al.[7]

Our model assumes no systematic secular trends in intake. This seems reasonable considering the short duration of the studies. Secular trends could bias estimates based on our model. In particular, the timing of biomarker measurements relative to FFQ administration would be a concern. Since FFQs inquire about past intake, one might expect biomarker measurements taken beforehand to correlate better with FFQ reports than those taken after. To check this, we compared such correlations for our studies, and found little difference between them. For example, the across-study mean correlation of FFQ energy with doubly-labeled water taken before the FFQ was 0.28 versus 0.27 for doubly-labeled water taken after the FFQ.

We also investigated calibration equations for predicting usual intakes, taking, as an example, equations that included the covariates available to us (age, BMI, race/ethnicity, and education). Choosing the covariates in a calibration model is complex. Covariates appearing in the disease model should also be included in the calibration model.[1] However, there is debate whether covariates in the calibration model should always be included in the disease model. BMI is particularly difficult because, while it is an important predictor of some intakes (e.g., energy, protein, and sodium), it is unclear whether it is a confounder that should be included in the disease model or a mediator that requires special methods.[5,37] Zheng et al.[38] discuss this issue and suggest that, alternatively, BMI be viewed as a marker for energy intake. Whichever view is taken, use of the time-varying model could better reflect the contribution of the self-report to the calibration equation.

The attenuation factors and correlations presented here are calculated assuming that the calibration models (see eAppendix, Supplemental Digital Content 1, models [A2] and [A3]; http://links.lww.com/EDE/A967) have zero coefficients for the covariates (Z). While this assumption could cause bias, we have found our attenuation factor estimates to be essentially unchanged when nonzero regression coefficients are allowed for these covariates. This happens because self-reported energy is only weakly associated with covariates, such as BMI, so the self-report coefficient estimates are little affected by including those covariates in the calibration model, leading to a close correspondence with the attenuation factors presented.

Our study has implications for nutritional epidemiology. First, it highlights the need to consider the period defining usual intake. Such consideration, in its broadest sense, involves

focusing on the period/s of life in which dietary intake is most expected to influence the health outcome of interest. This could impact on the choice of study population and the method of measuring intake. In most cohort studies of chronic disease, it is assumed that intake cumulated over adult life is the main interest, and thus the usual intake period will greatly exceed 1 year and may be closer to 30–50 years. However, we can assess attenuation factors and correlations only for periods as long as the validation studies that we conduct. As noted previously,[38] longer-term validation studies are needed to understand better longer-term variations in true intake, and the ability of our instruments to capture longer-term usual intake. The 6-year validation study conducted as part of the Nurses' Health Study[39] moves in this direction. Our study provides a framework for analysis of such longer-term studies.

Second, validation studies should pay attention to the timing of biomarker measurements relative to self-reports. Biomarker measurements and short-term self-reports should be spread over the targeted usual intake period, and the FFQ administered toward its end to compare with past diet as measured by the biomarkers. Furthermore, one should avoid taking self-reports and biomarker measurements of intakes on the same day (or otherwise close in time), as this leads to overestimating the correlation between self-report and usual intake under the fixed model.

Third, since attenuation factors and correlations for multiple 24HRs may be somewhat lower than previously estimated, there is even greater need to study combinations of multiple 24HRs with FFQs to achieve better accuracy.[40]

Allowing for variation in time of an exposure could be important in other areas of epidemiology. These include the study of physical activity, measurements of serum cholesterol and other biological precursors of heart disease, and secondhand smoke exposure. The study described here, together with Rosner et al.,[3] Keogh et al.,[4] Prentice and Huang,[5] and Zheng et al,[38] represents a general move to including the role of time in studying measurement error and its effects. This study should help to pave the way for dealing with future datasets that include longer-term longitudinal information on individuals' dietary intakes and life-course events.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Women's Health Initiative program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. Carroll's research was supported by a grant from the National Cancer Institute (U01-CA057030).

We thank the following people for their valuable contributions to this study: Dr. Alfonso Ang (University of California, Los Angeles), Dr. Sheila Bingham (deceased, Dunn Nutrition Unit, British Medical Research Council, Cambridge), Dr. Kevin Dodd (US National Cancer Institute), Dr. Nancy Potischman (National Cancer Institute), Dr. Dale Schoeller (University of Wisconsin), and Dr. Richard Troiano (National Cancer Institute). We also thank the WHI investigators and staff for their dedication. A full listing of Women's Health Initiative investigators can be found at http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf.

REFERENCES

1. Carroll, RJ.; Ruppert, D.; Stefanski, LA.; Crainiceanu, CM. Measurement Error in Nonlinear Models: A Modern Perspective. 2nd. Boca Raton, FL: Chapman and Hall; 2006.
2. Black AE, Cole TJ. 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. 2000; 54:386–394. [PubMed: 10822285]
3. Rosner B, Michels KB, Chen YH, Day NE. Measurement error correction for nutritional exposures with correlated measurement error: use of the method of triads in a longitudinal setting. Stat Med. 2008; 27:3466–3489. [PubMed: 18416440]
4. Keogh RH, White IR, Rodwell SA. Using surrogate biomarkers to improve measurement error models in nutritional epidemiology. Stat Med. 2013; 32:3838–3861. [PubMed: 23553407]
5. Prentice RL, Huang Y. Measurement error modeling and nutritional epidemiology association analyses. Can J Stat. 2011; 39:498–509. [PubMed: 23606784]
6. Freedman LS, Commins JM, Moler JE, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. Am J Epidemiol. 2014; 180:172–188. [PubMed: 24918187]
7. 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; 181:473–487. [PubMed: 25787264]
8. Kaaks R, Ferrari P, Ciampi A, et al. Uses and limitations of statistical accounting for random error correlations, in the validation of dietary questionnaire assessments. Pub Hlth Nutr. 2002; 5:969–976.
9. Bingham SA. Biomarkers in nutritional epidemiology. Public Health Nutr. 2002; 5:821–827. [PubMed: 12638591]
10. 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; 158:1–13. [PubMed: 12835280]
11. 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; 88:324–332. [PubMed: 18689367]
12. 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; 110:857–864. [PubMed: 20497774]
13. 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; 167:1247–1259. [PubMed: 18344516]
14. 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; 174:591–603. [PubMed: 21765003]
15. 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; 154:1089–1099. [PubMed: 11744511]
16. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol. 1985; 122:51–65. [PubMed: 4014201]
17. Patterson RE, Kristal AR, Tinker LF, et al. Measurement characteristics of the women's health initiative food frequency questionnaire. Ann Epidemiol. 1999; 9:178–187. [PubMed: 10192650]
18. Women's Health Initiative: Food Questionnaire. Available at: https://www.whi.org/studydoc/WHI%20Forms/F060%20v1.6.pdf.
19. Women's Health Initiative: WHI Procedure Manual. Section 10 Dietary Assessment. 2 Available at: https://www.whi.org/studydoc/WHI%20and%20ES1%20Manual%20of%20Operations/1993–2005%20WHI%20CT%20and%20OS/Vol%202,%2010%20-%20Dietary%20Assessment.pdf.
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; 126:348S–354S. [PubMed: 8558321]

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; 42:1276–1289. [PubMed: 4072961]
22. Mickelsen O, Makdani D, Gill JL, Frank RL. 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; 30:2033–2040. [PubMed: 930873]
23. Luft FC, Fineberg NS, Sloan RS. Estimating dietary sodium intake in individuals receiving a randomly fluctuating intake. Hypertension. 1982; 4:805–808. [PubMed: 7141607]
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; 134:1836–1843. [PubMed: 15226478]
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; 40:786–793. [PubMed: 6486085]
26. Kaaks R, Riboli E, van Staveren W. Calibration of dietary intake measurements in prospective cohort studies. Am J Epidemiol. 1995; 142:548–556. [PubMed: 7677134]
27. Kipnis V, Izmirlian G. The impact of categorization of continuous exposure measured with error (abstract). Am J Epidemiol. 2002; 155:S28.
28. Subar AF, Midthune D, Tasevska N, et al. Checking for completeness of 24-hour urine collection using para-amino benzoic acid (PABA) not necessary in the Observing Protein and Energy Nutrition (OPEN) Study. Eur J Clin Nutr. 2013; 67:863–867. [PubMed: 23486508]
29. Kipnis V, Subar AF, Midthune D, et al. The structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol. 2003; 158:14–21. [PubMed: 12835281]
30. SAS Institute Inc. Statistical Analysis System (SAS) software, Version 9.2. Cary, NC: SAS Institute Inc.;
31. Preis SR, Spiegelman D, Zhao BB, et al. Application of a repeat-measure biomarker measurement error model to two validation studies: examination of the effect of within-person variation in biomarker measurements. Am J Epidemiol. 2011; 173:683–694. [PubMed: 21343245]
32. Dodd KW, Midthune D, Kipnis V. Application of a repeat-measure biomarker measurement error model to two validation studies: examination of the effect of within-person variation in biomarker measurements (letter). Am J Epidemiol. 2012; 175:84–85. [PubMed: 22088499]
33. 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. [PubMed: 9094927]
34. Bingham SA, Day NE. Using biochemical markers to assess the validity of prospective dietary assessment methods and the effect of energy adjustment. Am J Clin Nutr. 1997; 65(4 Suppl): 1130S–1137S. [PubMed: 9094909]
35. Freisling H, van Bakel MM, Biessy C, et al. Dietary reporting errors on 24 h recalls and dietary questionnaires are associated with BMI across six European countries as evaluated with recovery biomarkers for protein and potassium intake. Br J Nutr. 2012; 107:910–920. [PubMed: 21791145]
36. Turban S, Thompson CB, Parekh RS, Appel LJ. Effects of sodium intake and diet on racial differences in urinary potassium excretion: results from the Dietary Approaches to Stop Hypertension (DASH)-Sodium trial. Am J Kidney Dis. 2013; 61:88–95. [PubMed: 23157935]
37. 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; 174:1238–1245. [PubMed: 22047826]
38. Zheng C, Beresford SA, Van Horn L, et al. Simultaneous association of total energy consumption and activity-related energy expenditure with risks of cardiovascular disease, cancer, and diabetes among postmenopausal women. Am J Epidemiol. 2014; 180:526–535. [PubMed: 25016533]
39. Willett, W.; Lenart, E. Reproducibility and validity of food frequency questionnaires. In: Willett, W., editor. Nutritional Epidemiology. 3rd. New York, NY: Oxford University Press; 2013. p. 96-141.
40. Carroll RJ, Midthune D, Subar AF, et al. Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. Am J Epidemiol. 2012; 175:340–347. [PubMed: 22273536]

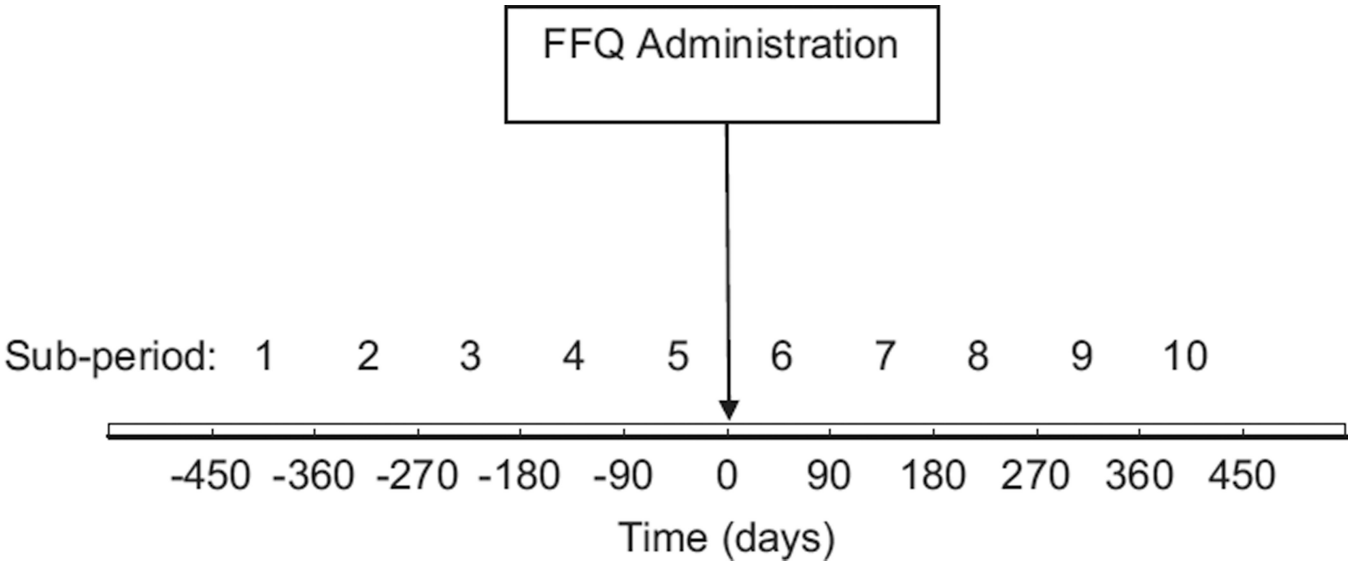

FIGURE 1.
Subdivision of time period into subperiods.

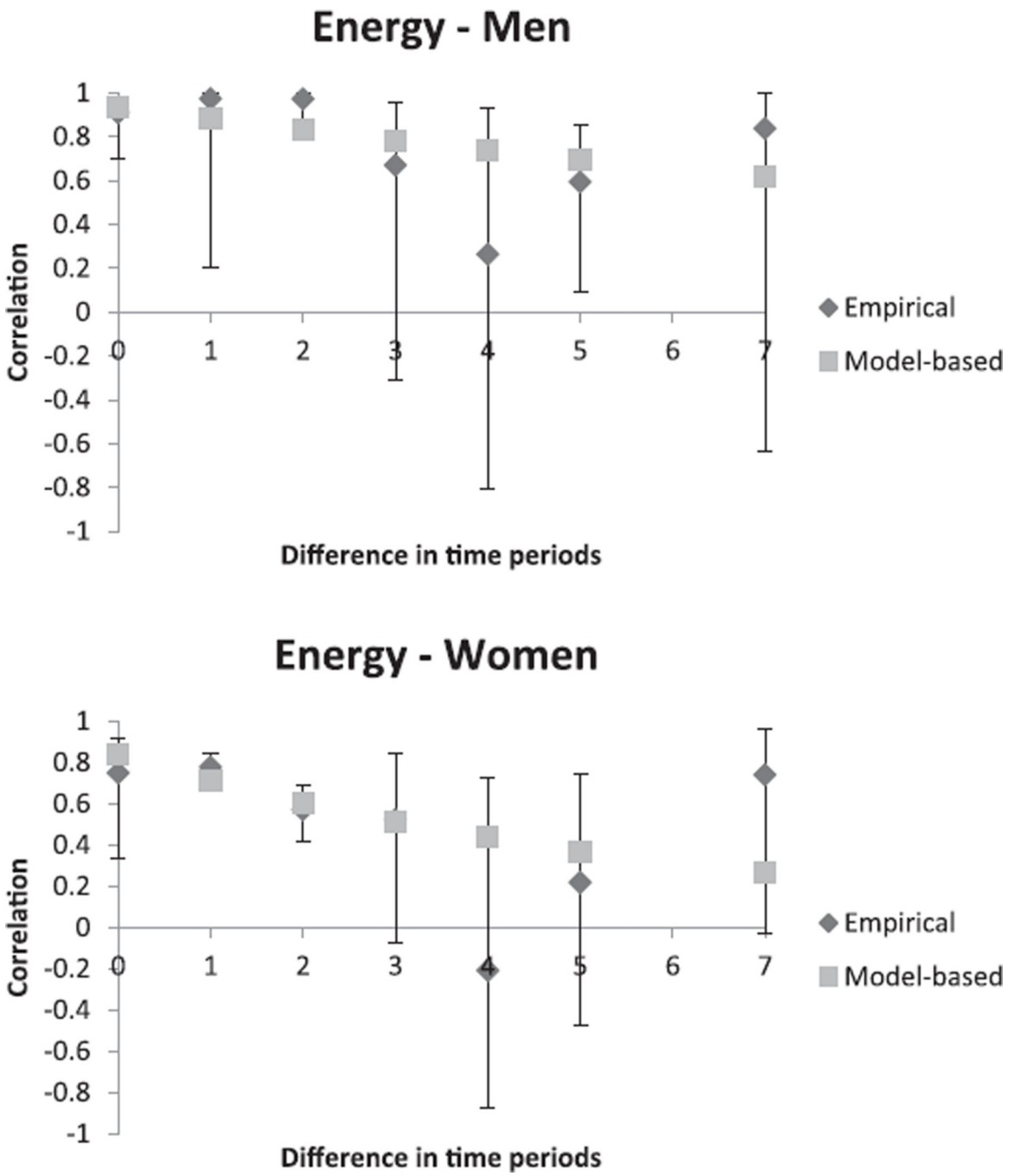

FIGURE 2.

Empirical and time-varying usual intake model-based estimates of correlation between biomarker measurements for energy intake according to the time difference between them (measured in 3-month time periods). The *error bars* shown are 95% confidence intervals for the empirical correlations.

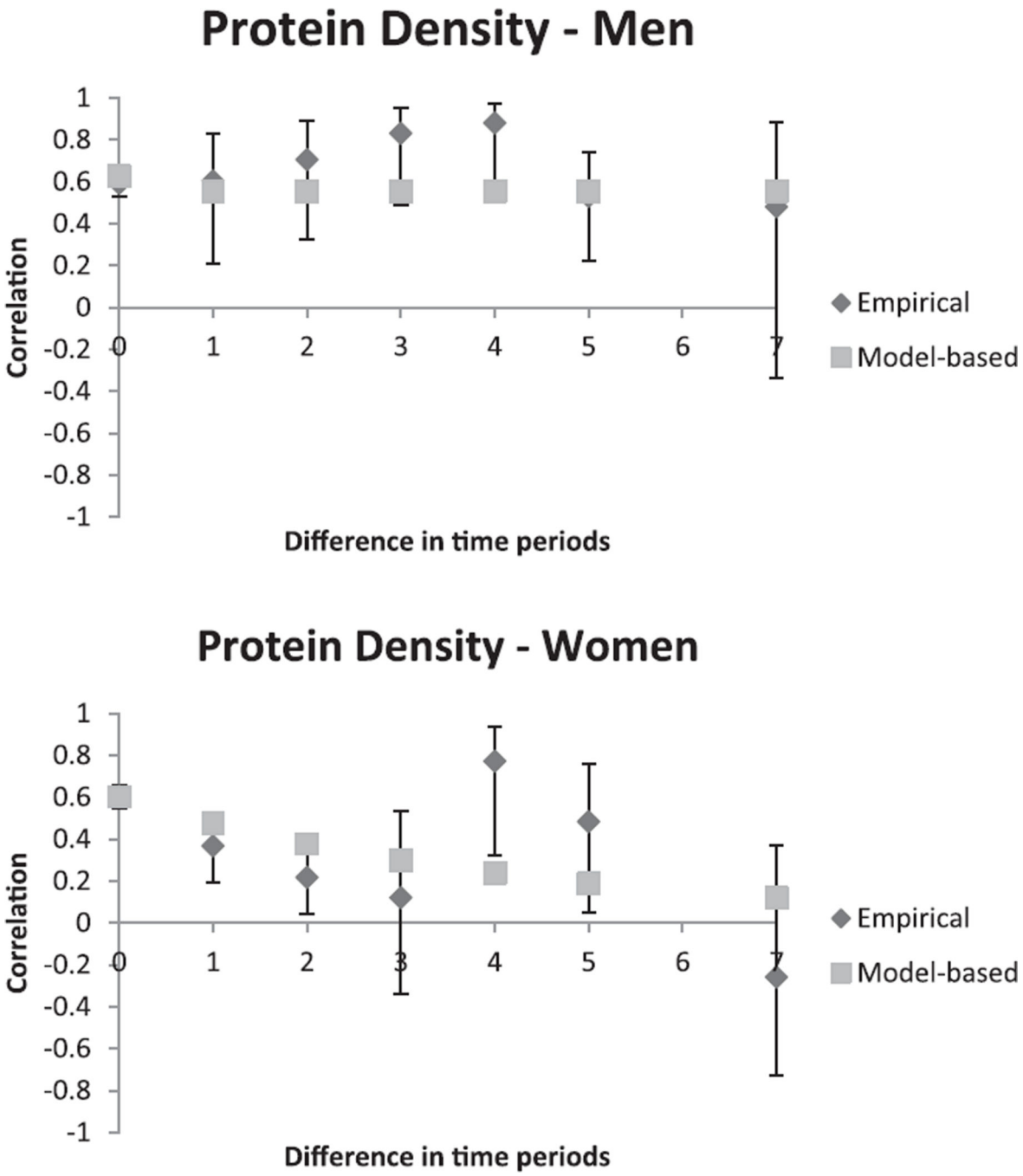

FIGURE 3.

Empirical and time-varying usual intake model-based estimates of correlation between biomarker measurements for protein density intake according to the time difference between them (measured in 3-month time periods). The *error bars* shown are 95% confidence intervals for the empirical correlations.

TABLE 1

Correlation Structure for Each Dietary Component and Sex

Sex	Dietary Component	Correlation Structure	Parameter ρ (SE)	Intake Variance Ratio: Day vs. Sub-period (SE)	AIC[a]	AIC for Fixed Usual Intake Model
Men	Energy	AR(1)	0.94 (0.02)	_[b]	908	916
	Protein	CS	0.87 (0.07)	1.30 (0.09)	2,046	2,070
	Potassium	AR(1)	0.87 (0.06)	1.47 (0.12)	2,352	2,404
	Sodium	AR(1)	0.92 (0.05)	1.46 (0.18)	3,417	3,435
	Protein density	CS	0.89 (0.08)	1.44 (0.13)	−66	−34
	Potassium density	AR (1)	0.92 (0.04)	1.43 (0.10)	907	969
	Sodium density	AR(1)	0.91 (0.05)	1.50 (0.20)	1,323	1,341
Women	Energy	AR(1)	0.85 (0.03)	-	2,003	2,025
	Protein	AR(1)	0.83 (0.05)	1.50 (0.11)	5,525	5,603
	Potassium	AR(1)	0.94 (0.02)	1.42 (0.11)	5,947	5,993
	Sodium	CS	0.55 (0.10)	1.39 (0.15)	8,333	8,387
	Protein density	AR(1)	0.79 (0.05)	1.50 (0.16)	682	748
	Potassium density	AR(1)	0.91 (0.03)	1.31 (0.11)	2,821	2,863
	Sodium density	CS	0.47 (0.12)	1.43 (0.20)	3,295	3,334

[a] Akaike information criterion = −2 × log likelihood + 2×(no. parameters in model) (smaller is better).

[b] Could not be estimated, because the biomarker (doubly labeled water) does not assess a single day's intake.

AIC indicates Akaike information criterion; AR(1), autoregressive model of order 1; CS, compound symmetric model; SE, standard error.

TABLE 2

Estimates of Overall[a] Attenuation Factors and Correlations with Truth for the Fixed and Time-varying Usual Intake Models Without Covariates (Standard Errors in Parentheses): Energy, Protein, and Protein Density

			Attenuation Factor		Correlation with Truth	
Dietary Component	**Instrument**	**Sex**	**Fixed Model**	**Time-varying Model**	**Fixed Model**	**Time-varying Model**
Energy	FFQ[b]	M	0.044 (0.020)	0.046 (0.019)	0.092 (0.046)	0.095 (0.045)
		F	0.072 (0.011)	0.074 (0.011)	0.217 (0.032)	0.207 (0.032)
	1 × 24HR[c]	M	0.103 (0.015)	0.100 (0.014)	0.276 (0.035)	0.257 (0.032)
		F	0.084 (0.010)	0.072 (0.010)	0.219 (0.027)	0.178 (0.025)
	4 × 24HR[d]	M	0.186 (0.026)	0.180 (0.025)	0.374 (0.046)	0.349 (0.042)
		F	0.165 (0.020)	0.144 (0.019)	0.307 (0.037)	0.251 (0.034)
Protein	FFQ	M	0.168 (0.026)	0.163 (0.026)	0.295 (0.045)	0.298 (0.047)
		F	0.169 (0.016)	0.180 (0.017)	0.329 (0.030)	0.336 (0.032)
	1 × 24HR	M	0.207 (0.019)	0.182 (0.019)	0.413 (0.030)	0.355 (0.032)
		F	0.222 (0.015)	0.201 (0.015)	0.397 (0.023)	0.340 (0.025)
	4 × 24HR	M	0.423 (0.035)	0.344 (0.033)	0.594 (0.037)	0.491 (0.047)
		F	0.463 (0.029)	0.376 (0.026)	0.580 (0.030)	0.476 (0.031)
Protein density	FFQ	M	0.423 (0.052)	0.420 (0.052)	0.390 (0.045)	0.396 (0.048)
		F	0.396 (0.033)	0.439 (0.036)	0.422 (0.035)	0.450 (0.039)
	1 × 24HR	M	0.266 (0.026)	0.240 (0.025)	0.365 (0.031)	0.328 (0.032)
		F	0.235 (0.020)	0.184 (0.021)	0.356 (0.027)	0.263 (0.031)
	4 × 24HR	M	0.612 (0.055)	0.481 (0.048)	0.554 (0.042)	0.483 (0.042)
		F	0.534 (0.042)	0.384 (0.041)	0.537 (0.037)	0.382 (0.043)

[a] Weighted average over studies, where weights are the inverse of the variance.

[b] Food frequency questionnaire.

[c] Single 24-hour recall.

[d] Average of four repeats of a 24-hour recall.

TABLE 3

Estimates of Overall[a] Attenuation Factors and Correlations with Truth for the Fixed and Time-varying Usual Intake Models Without Covariates (Standard Errors in Parentheses): Potassium, Sodium, and Their Densities

			Attenuation Factor		Correlation with Truth	
Dietary Component	**Instrument**	**Sex**	**Fixed Model**	**Time-varying Model**	**Fixed Model**	**Time-varying Model**
Potassium	FFQ[b]	M	0.302 (0.033)	0.321 (0.035)	0.452 (0.044)	0.484 (0.053)
		F	0.248 (0.022)	0.257 (0.022)	0.368 (0.031)	0.371 (0.031)
	1 × 24HR[c]	M	0.282 (0.023)	0.252 (0.024)	0.502 (0.030)	0.417 (0.035)
		F	0.327 (0.018)	0.311 (0.019)	0.484 (0.023)	0.451 (0.025)
	4 × 24HR[d]	M	0.547 (0.040)	0.398 (0.037)	0.685 (0.036)	0.541 (0.041)
		F	0.573 (0.030)	0.492 (0.031)	0.642 (0.028)	0.573 (0.030)
Potassium density	FFQ	M	0.525 (0.054)	0.528 (0.055)	0.464 (0.042)	0.485 (0.047)
		F	0.546 (0.035)	0.568 (0.037)	0.509 (0.030)	0.517 (0.032)
	1 × 24HR	M	0.424 (0.032)	0.395 (0.033)	0.507 (0.029)	0.445 (0.033)
		F	0.361 (0.024)	0.344 (0.024)	0.444 (0.025)	0.425 (0.026)
	4 × 24HR	M	0.813 (0.055)	0.646 (0.052)	0.689 (0.034)	0.573 (0.038)
		F	0.658 (0.041)	0.558 (0.041)	0.604 90.031)	0.560 (0.033)
Sodium	FFQ	M	0.113 (0.035)	0.121 (0.036)	0.181 (0.056)	0.195 (0.060)
		F	0.085 (0.023)	0.089 (0.023)	0.138 (0.039)	0.154 (0.043)
	1 × 24HR	M	0.186 (0.024)	0.168 (0.024)	0.359 (0.038)	0.321 (0.040)
		F	0.138 (0.018)	0.096 (0.016)	0.248 (0.030)	0.177 (0.029)
	4 × 24HR	M	0.393 (0.047)	0.326 (0.043)	0.524 (0.051)	0.449 (0.051)
		F	0.321 (0.039)	0.216 (0.032)	0.381 (0.043)	0.268 (0.042)
Sodium density	FFQ	M	0.368 (0.072)	0.396 (0.075)	0.308 (0.063)	0.348 (0.068)
		F	0.343 (0.049)	0.346 (0.051)	0.277 (0.047)	0.302 (0.052)
	1 × 24HR	M	0.249 (0.031)	0.228 (0.030)	0.347 (0.042)	0.319 (0.042)
		F	0.174 (0.021)	0.099 (0.019)	0.255 (0.033)	0.149 (0.030)
	4 × 24HR	M	0.593 (0.070)	0.478 (0.061)	0.544 (0.063)	0.483 (0.057)
		F	0.459 (0.055)	0.259 (0.045)	0.416 (0.052)	0.245 (0.047)

[a]Weighted average over studies, where weights are the inverse of the variance.

[b] Food frequency questionnaire.

[c] Single 24-hour recall.

[d] Average of four repeats of a 24-hour recall.