Research Article

Cancer Epidemiology, Biomarkers & Prevention

Use of a Urinary Sugars Biomarker to Assess Measurement Error in Self-Reported Sugars Intake in the Nutrition and Physical Activity Assessment Study (NPAAS)

Natasha Tasevska[1,2], Douglas Midthune[1], Lesley F. Tinker[3], Nancy Potischman[1], Johanna W. Lampe[3], Marian L. Neuhouser[3], Jeannette M. Beasley[4], Linda Van Horn[5], Ross L. Prentice[3], and Victor Kipnis[1]

Abstract

Background: Measurement error in self-reported sugars intake may be obscuring the association between sugars and cancer risk in nutritional epidemiologic studies.

Methods: We used 24-hour urinary sucrose and fructose as a predictive biomarker for total sugars, to assess measurement error in self-reported sugars intake. The Nutrition and Physical Activity Assessment Study (NPAAS) is a biomarker study within the Women's Health Initiative (WHI) Observational Study that includes 450 postmenopausal women ages 60 to 91 years. Food Frequency Questionnaires (FFQ), four-day food records (4DFR), and three 24-hour dietary recalls (24HRs) were collected along with sugars and energy dietary biomarkers.

Results: Using the biomarker, we found self-reported sugars to be substantially and roughly equally misreported across the FFQ, 4DFR, and 24HR. All instruments were associated with considerable intake- and person-specific bias. Three 24HRs would provide the least attenuated risk estimate for sugars (attenuation factor, AF = 0.57), followed by FFQ (AF = 0.48) and 4DFR (AF = 0.32), in studies of energy-adjusted sugars and disease risk. In calibration models, self-reports explained little variation in true intake (5%–6% for absolute sugars and 7%–18% for sugars density). Adding participants' characteristics somewhat improved the percentage variation explained (16%–18% for absolute sugars and 29%–40% for sugars density).

Conclusions: None of the self-report instruments provided a good estimate of sugars intake, although overall 24HRs seemed to perform the best.

Impact: Assuming the calibrated sugars biomarker is unbiased, this analysis suggests that measuring the biomarker in a subsample of the study population for calibration purposes may be necessary for obtaining unbiased risk estimates in cancer association studies. *Cancer Epidemiol Biomarkers Prev; 23(12); 2874–83. ©2014 AACR.*

Introduction

The associations between sugars consumption and cancer have long been studied yet are difficult to demonstrate in an epidemiologic context (1–3). One of the major limitations in nutritional epidemiologic analyses is unreliability of self-reported dietary data, which can lead to severely distorted disease association estimates and reduced statistical power to detect an association (4).

Food frequency questionnaires (FFQ) are the most widely used dietary assessment instruments in population studies, despite known large measurement errors both random and systematic (5, 6). Other self-reports such as 24-hour dietary recalls (24HRs; ref. 7) and food records (FR; ref. 8), although generally considered to have better measurement error properties than FFQs, they too are affected by random and systematic measurement errors (7, 9, 10). These findings have prompted researchers to incorporate reference dietary biomarkers as objective measures of intake in subsamples of their cohorts to assess and correct for measurement errors in self-reported intake. So far, 4 classes of biomarkers have been described (11). *Recovery* biomarkers are based on a known recovered proportion of intake over certain period of time. Following transformation, recovery biomarkers adhere to a classical measurement error model and generate unbiased estimates of intakes (12), which can be used to assess measurement errors in self-report instruments (7, 10) or to develop calibration equations for calibrating (i.e.,

[1]National Cancer Institute, Bethesda, Maryland. [2]School of Nutrition and Health Promotion, Arizona State University, Phoenix, Arizona. [3]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington. [4]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York. [5]Department of Preventive Medicine, Northwestern University, Chicago, Illinois.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Natasha Tasevska, School of Nutrition and Health Promotion, Arizona State University, 500 North 3rd Street, Phoenix, AZ 85004. Phone: 602-827-2485; Fax: 602-827-2253; E-mail: natasha.tasevska@asu.edu

doi: 10.1158/1055-9965.EPI-14-0594

correcting) self-reported intake to be applied in diet–disease risk models of association studies (13). *Concentration* biomarkers provide correlate rather than a direct measure of intake (14), yet when combined with self-reported intake were shown to improve reliability of risk estimates and to increase the statistical power to detect an association (15). *Replacement* biomarkers replace estimates of intake for nutrients or compounds difficult to measure or with no food composition data available and depending on their innate characteristics may be used as recovery (e.g., 24-hour urinary sodium) or more commonly as concentration biomarkers (e.g., serum phytoestrogens). *Predictive* biomarkers, the most recently described class of biomarkers, predict intake after being calibrated to account for certain level of bias, estimable from a feeding study and assumed to be stable across populations (16). Following calibration, similar to recovery biomarkers, they, too, can be used as reference instruments.

Recently, on the basis of findings from 2 controlled feeding studies, 24-hour urinary sucrose plus fructose was suggested as a predictive biomarker for total sugars intake (17). Although predictive biomarkers exhibit more complex relationship with true intake than recovery biomarkers, this relationship is assumed stable, thereby distinguishing predictive biomarkers from less specific concentration biomarkers (16). On the basis of a novel measurement error model for predictive biomarkers (16), the sugars biomarker was calibrated for use as a reference instrument in biomarker studies of dietary measurement errors or disease association cohort studies having available 24-hour urine collections in a subsample of participants.

The Nutrition and Physical Activity Assessment Study (NPAAS) is a biomarker study involving a subset of 450 participants in the Women's Health Initiative (WHI) Observational Study. Prentice and colleagues (10) compared FFQ, 24HRs, and 4DFR energy and protein intake in NPAAS against respective recovery biomarkers, doubly labeled water (DLW), and 24-h urinary nitrogen. We now compare estimates of total sugars intake from these 3 self-report instruments against the 24-hour urinary sugars biomarker. After calibrating the biomarker, we use it as a reference measure to evaluate the measurement error structure of self-reported sugars and to estimate attenuation factors and correlations with true intake, 2 important parameters that determine how well each instrument will be able to detect and estimate disease risks associated with total sugars intake. We also develop calibration equations that predict total sugars intake, based on the objective predictive sugars biomarker given self-report and other covariates that could be applied in future WHI association studies for more reliable disease risk estimation.

Materials and Methods

Participants

The NPAAS is an ancillary study to the WHI Observational Study, a prospective study of 93,676 postmenopausal women ages 50 to 79 years, enrolled during 1994–1998 at 40 U.S. clinical centers (18, 19). Between 2007 and 2009, 450 participants ages 60 to 91 years at the time of the NPAAS were recruited from 9 WHI centers. The NPAAS design has been previously described in detail (10). To ensure adequate statistical power to assess the effect of racial/ethnic groups, body mass index (BMI), and age, recruitment was conducted to oversample African-American and Hispanic women, women with BMI < 18.5 and $\geq$30 kg/m^2, and those who were 50 to 59 years at WHI enrollment, respectively. Among women who were screened, 20.6% were eligible and willing to participate. The study was approved by the Institutional Review Boards of all participating institutions.

Study design

Participants visited local WHI centers twice. At visit 1, they provided informed consent, had their body weight and height measured, and completed the FFQ and the WHI Personal Habit Questionnaire (PHQ) assessing participants' recreational physical activity (20). Information on participants' demographic characteristics was available from the WHI database. During visit 1, participants received training on how to keep a 4DFR and started the 2-week DLW protocol for measurement of energy expenditure, which was concluded at visit 2 (10). In the 2 weeks between visits 1 and 2, participants completed the 4DFR and collected 24-hour urine on the day before visit 2. Three 24HRs were administered following visit 2. Approximately 20% of the NPAAS participants ($n = 88$) completed the "reliability study," which involved repeating the whole study protocol 6 months after the baseline study.

Dietary assessment

Participants' diet over the previous 3 months was assessed at NPAAS baseline using the WHI self-administered semiquantitative FFQ (21), designed to assess diet in a multiethnic and geographically diverse population, and inquired about 122 foods or food groups, including 19 adjustment questions and 4 summary questions. The FFQs were reviewed by study staff, checking for missing responses in the presence of the participants. As part of the 4DFR training, participants viewed 25-minute instructional video and received instructions along with a serving size booklet containing photographs and measuring devices. They completed the 4DFR between visits 1 and 2 over 4 alternate days, including 1 weekend day. Upon return, the 4DFR was reviewed in the presence of participants. For the 3 24HRs, the first was administered 1 to 3 weeks after visit 2 and the other 2 approximately monthly thereafter by a certified interviewer over the phone (70% weekday and 30% weekend) with computer-assisted data entry using the NDSR software (Nutrition Data System for Research; Nutrition Coordinating Center, University of Minnesota, Minneapolis, MN) based on the USDA multipass method (22). The NDSR version 2005 was used to calculate total sugars intake (g/d), a sum of sucrose, fructose, glucose, lactose, galactose, and maltose, and total sugars density (g/1,000 kcal) for the 3 instruments. On the

basis of information provided on the 4DFR, we calculated percentage meals eaten at home.

Urinary sugars biomarkers

The day before visit 2, participants collected a 24-hour urine sample. They were asked to record any missed voids or spillage during the collection. To assess urine completeness, participants took three 80-mg tablets of para-aminobenzoic acid (PABA) on the day of urine collection, one with each meal (23). In this article, we analyzed 384 participants from the primary and 78 from the reliability study, who provided urine samples, had <2 missed voids or <240 mL of missed/spilled urine, and had sufficient specimen available for laboratory analysis, without excluding urine samples based on PABA (see Supplementary Material for results with PABA exclusions).

Sucrose and fructose in urine were quantified by a colorimetric enzymatic assay (sucrose/D-glucose/D-Fructose; Boehringer Mannheim, R-Biopharm, Roche), modified to be run on a microplate reader in 96-well plates. Samples were run in duplicate, and glucose, sucrose, and fructose standards (5, 10, 20, 30, 50, and 100 mg/L) and quality control urine samples were included in each run. When percentage of coefficient of variance (CV) for duplicate samples was >15%, samples were re-analyzed. When applicable, urine samples were diluted and re-analyzed to obtain values within the range of the calibration curve. The level of detection (LOD) was 1 mg/L. The intra- and interassay CV for the method were <18% and <25% for concentrations <5 mg/L and <10% and <20% for concentrations $\geq$5 mg/L, respectively, for both sucrose and fructose. Urinary excretion of sucrose and fructose (mg/d) was calculated on the basis of 24-hour urine volume (mL/d). Nearly 139 sucrose and 7 fructose values were below LOD, in which case a value of 0.5 mg/L was imputed.

Biomarker for energy intake

We used total energy expenditure, assessed by DLW as previously described (10). It has been shown that in weight-stable individuals, 2 weeks of total energy consumption can be objectively estimated by this approach (24, 25). Approximately 6.5% of the samples was excluded from the analysis; half due to low tracer enrichments or lack of equilibration, whereas the others due to dilution space or other external reproducibility issues. DLW-based estimates of energy expenditure (kcal/d) were used to express the biomarker-based total sugars density (g/1,000 kcal).

Statistical methods

Calibration of the urinary sugars biomarker. We used the calibration equation for the total sugars biomarker developed by Tasevska and colleagues (16), based on data from a UK-based 30-day feeding study with 13 subjects ages 23 to 66 years, consuming their usual diets and collecting 24-hour urines for 30 days under highly controlled conditions (17).

$$M^*_{ij} = M_{ij} - 1.67 - 0.02 \times S_i + 0.71 \times A_i, \quad (1)$$

where (M_{ij}) is log-transformed sugars biomarker, S_i is an indicator variable that equals 0 for men and 1 for women, and A_i is log-transformed age. The calibrated biomarker M^*_{ij} satisfies the following measurement errors model

$$M^*_{ij} = T_i + u_{Mi} + \varepsilon_{Mij}, \quad (2)$$

where T_i is log-transformed true usual intake of total sugars, u_{Mi} is a person-specific bias (random-effect) having mean zero and variance $\sigma^2_{u_M}$, and ε_{Mij} is within-person random error. To use the calibrated biomarker as a reference measure in the measurement errors model for self-reported sugars intake, we would need to know the variance $\sigma^2_{u_M}$ or some related value such as the ratio of the variance of u_{Mi} to the variance of $T_i + u_{Mi}$, calculated as

$$\kappa = \frac{\sigma^2_{u_M}}{\sigma^2_T + \sigma^2_{u_M}}. \quad (3)$$

In our analysis, we used the value $\kappa = 0.218$, estimated from the feeding study (16).

Measurement errors model for self-reported sugars intake. The measurement errors model for self-reported sugars intake has been described in detail (16). Briefly, for individual i, let Q_{ij}, F_{ij}, and R_{ij} denote log-transformed reported intake on the j^{th} application of the FFQ, 4DFR, and 24HR, respectively. Let X_{Qi}, X_{Fi}, and X_{Ri} be vectors of covariates that affect the relationship between reported and true usual intake T_i. The measurement errors model is:

$$\begin{aligned} Q_{ij} &= \beta_{Q0} + \beta_{QT} T_i + \beta_{QX} X_{Qi} + u_{Qi} + \varepsilon_{Qij}, \\ F_{ij} &= \beta_{F0} + \beta_{FT} T_i + \beta_{FX} X_{Fi} + u_{Fi} + \varepsilon_{Fij}, \\ R_{ij} &= \beta_{R0} + \beta_{RT} T_i + \beta_{RX} X_{Ri} + u_{Ri} + \varepsilon_{Rij}, \\ M^*_{ij} &= T_i + u_{Mi} + \varepsilon_{Mij}, \end{aligned} \quad (4)$$

where M^*_{ij} is the calibrated biomarker described in the previous section; u_{Qi}, u_{Fi}, u_{Ri}, and u_{Mi} are person-specific biases with zero means and variances $\sigma^2_{u_Q}, \sigma^2_{u_F}, \sigma^2_{u_R}$ and $\sigma^2_{u_M}$, respectively, and e_{Qij}, e_{Fij}, e_{Rij}, and e_{Mij}, are within-person random errors with zero means and variances $\sigma^2_{\varepsilon_Q}, \sigma^2_{\varepsilon_F}, \sigma^2_{\varepsilon_R}$ and $\sigma^2_{\varepsilon_M}$, respectively.

The main assumption is that the person-specific biases of the FFQ, 4DFR, and 24HR (u_{Qi}, u_{Fi}, u_{Ri}), although possibly correlated with each other, are independent of the person-specific bias of the biomarker (u_{Mi}). The parameters in model (4) were estimated using maximum likelihood. Log age and log BMI were included as components of the covariate vector $X_{Qi} = X_{Fi} = X_{Ri}$, as it is believed these variables may affect the relationship between self-report and true intake (26).

The estimated parameters in the measurement errors model can be used to estimate 2 characteristics that are important for evaluating the effect of measurement error

in self-reported intake on the estimation of diet–disease associations; the attenuation factor (AF) is the multiplicative bias in the estimated association due to using self-reported rather than true intake, whereas the Pearson correlation coefficient (CC) between self-reported and true intake is related to the loss of power to detect the association due to measurement errors. Formulas for calculating attenuation factor and correlation coefficient are given in the study by Tasevska and colleagues (16).

We also developed calibration equations for predicting true intake by regressing M^*_{ij} on Q_{ij}, F_{ij}, or R_{ij}, and other covariates that may be related to self-reported sugars and/or true intake. Potential covariates included age, BMI, physical activity (metabolic equivalents/wk), smoking status, dietary supplement use, race, education, annual income, and percentage of meals eaten at home. The latter covariate was considered because low percentage of meals eaten at home was associated with energy underreporting in NPAAS (27). Metabolic equivalents values for physical activity reported on the PHQ were generated as previously described (28). The calibration model included log age and log BMI. Backward selection ($P = 0.1$) was used to select other covariates that were significant predictors of log true intake. We reported R^2 adjusted for within-subject error in the biomarker, using repeated biomarker measurements from the reliability study ($n = 88$). We also reported unadjusted R^2 because of the possibility of correlated measurement errors for the biomarker collected 6 months apart.

We performed sensitivity analyses to examine the effects of excluding incomplete 24-hour urine collections based on PABA (Supplementary Material Section A) and of calibrating the biomarker using a measurement errors model that did not include age (Supplementary Material Section B).

Results

Table 1 reports baseline characteristics for the participants. Geometric means and 95% confidence intervals (CI) for self-reported and biomarker-based total sugars intake (g/d) and total sugars density (g/1,000 kcal) are presented in Table 2. Compared with biomarker-based estimates, self-reported total sugars appeared considerably misreported on all 3 instruments (Figs. 1–3), with geometric means approximately half those of the biomarker. The underreporting of total sugars density was less striking, with the ratio of self-reported to biomarker-based means ranging from 0.6 to 0.7. Geometric means of daily sucrose and fructose urinary excretion were 10.7 and 20.0 mg/d in the primary and 12.5 and 22.3 mg/d in the reliability study, respectively. Sugars excretion, particularly sucrose excretion, was highly variable between participants, with almost a third of participants having urinary sucrose concentration below LOD.

Measurement error model parameters for the FFQ, 4DFR, and 24HR are reported in Table 3. For total sugars (g/d), the slope in the regression of reported on true intake (β_{Q1}, β_{F1}, or β_{R1}) was approximately 0.2 for all instruments. This slope measures the extent to which bias in self-report is related to true intake: $\beta_{Q1} = 1$ indicates no intake-related bias, whereas $\beta_{Q1} < 1$ indicates the extent to which participants tend to underreport high and overreport low intakes (a "flattened slope phenomenon"). A value of 0.2

Table 1. Demographics and lifestyle characteristics among NPAAS participants (primary study, $n = 450$; reliability study, $n = 88$) collected at WHI (1994–1998) or NPAAS (2007–2009) baseline

Characteristic	*n* (%)
Age,[a] y	
60–64	79 (17.5)
65–69	165 (36.7)
70–74	110 (24.4)
75–79	59 (13.1)
≥80	37 (8.2)
BMI[a]	
<18.5 kg/m^2	9 (2)
18.5–24.9 kg/m^2	143 (31.8)
25.0–29.9 kg/m^2	125 (27.8)
≥30 kg/m^2	173 (38.4)
Race/ethnicity	
Non-Hispanic white	288 (64)
African American	84 (18.6)
Hispanic	64 (14.2)
Asian/Pacific Islander	8 (1.8)
Other	6 (1.3)
Annual income	
<$20,000	43 (9.9)
$20,000–$49,999	176 (40.6)
≥$50,000	215 (49.5)
Education level	
None to some high school	16 (3.6)
High school graduate	48 (10.7)
Some college	157 (35.1)
College graduate or higher	226 (50.6)
Smoking status	
Current	21 (4.7)
Former or never	424 (95.3)
Recreational physical activity (metabolic equivalents/wk)	
<5	111 (25.0)
5–19.9	190 (42.8)
≥20.0	143 (32.2)
Meals eaten at home[b] (%)	
<75.0	103 (22.9)
75.0–94.9	264 (58.7)
≥95.0	83 (18.4)

[a]Age and BMI were collected at NPAAS baseline, whereas all other variables were obtained at WHI baseline.
[b]Measured by 4DFR.

Table 2. Geometric means with 95% CI and CVs for total sugars intake and total sugars density assessed by FFQ, 24HR, 4DFR, and urinary sugars biomarker among NPAAS (2007–2009) participants (primary study, $n = 450$; reliability study, $n = 88$)

	Total sugars intake, g/d			Total sugars density, g/1,000 kcal		
Instrument	*n*	Geometric mean (95% CI)	CV	*n*	Geometric mean (95% CI)	CV
FFQ[a]	450	82.4 (78.7–86.4)	0.54	450	56.6 (55.0–58.3)	0.32
FFQ[b]	88	82.8 (75.1–91.3)	0.50	88	55.1 (51.4–59.1)	0.34
4DFR[a]	450	86.5 (83.3–89.8)	0.42	450	53.5 (51.9–55.1)	0.33
4DFR[b]	88	83.8 (77.4–90.7)	0.39	88	52.0 (49.0–55.2)	0.29
First 24HR[a]	398	80.1 (75.8–84.6)	0.60	398	52.3 (50.1–54.6)	0.46
Second 24HR[a]	425	80.2 (76.0–84.6)	0.61	425	54.0 (51.8–56.3)	0.46
Third 24HR[a]	421	82.0 (77.9–86.3)	0.57	421	53.5 (51.4–55.6)	0.43
24HR[a] (3-day avg.)	440	85.1 (81.8–88.5)	0.44	440	54.8 (53.2–56.4)	0.32
First 24HR[b]	86	78.7 (69.9–88.7)	0.61	86	52.7 (48.2–57.8)	0.45
Second 24HR[b]	83	81.5 (71.2–93.4)	0.70	83	51.4 (45.8–57.8)	0.58
Third 24HR[b]	82	72.3 (62.7–83.4)	0.74	82	48.8 (43.4–55.0)	0.59
24HR[b] (3-day avg.)	87	82.4 (75.1–90.5)	0.47	87	53.1 (49.4–57.2)	0.36
Biomarker[a,c]	384	158.8 (145.6–173.1)	1.06	351	78.3[d] (71.6–85.6)	1.03
Biomarker[b,c]	78	173.9 (142.9–211.6)	1.09	73	82.9[d] (67.3–102.2)	1.14

[a]Collected in the primary study.
[b]Collected in the reliability study.
[c]Calibrated using the measurement error parameters generated from the feeding study (17).
[d]Energy intake estimated by DLW measurement of total energy expenditure.

denotes large intake-related bias. Energy adjustment using total sugars density did not improve the intake-related bias for any of the instruments.

The variance of the person-specific bias in total sugars intake was much larger for FFQ (0.15) than for 4DFR (0.07) or 24HR (0.07). Energy adjustment led to smaller person-specific biases for all instruments (0.04–0.06). Person-specific biases in both 4DFR and 24HR were highly correlated with person-specific bias in FFQ [~0.6 for total sugars and 0.70 (24HR) and 0.84 (4DFR) for density]. The variance of within-person error in 24HR was 3 times as large as that in FFQ or 4DFR and remained the greatest after energy adjustment.

The attenuation factor measures the level of attenuation (underestimation) of disease risk due to measurement errors in self-reported intake; AF = 1 indicates no

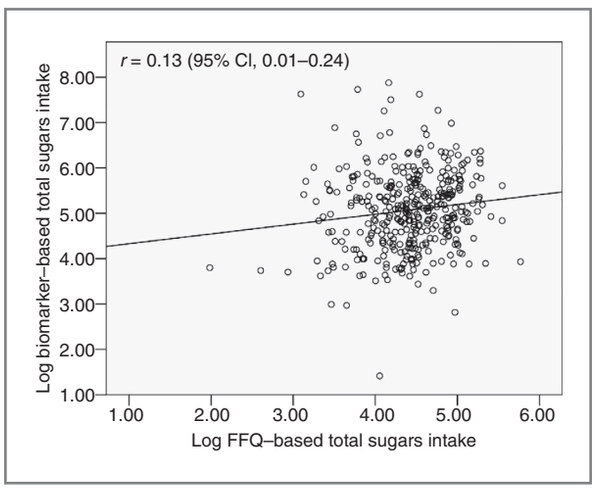

Figure 1. Association between log biomarker–based and log FFQ–based total sugars intake among NPAAS participants.

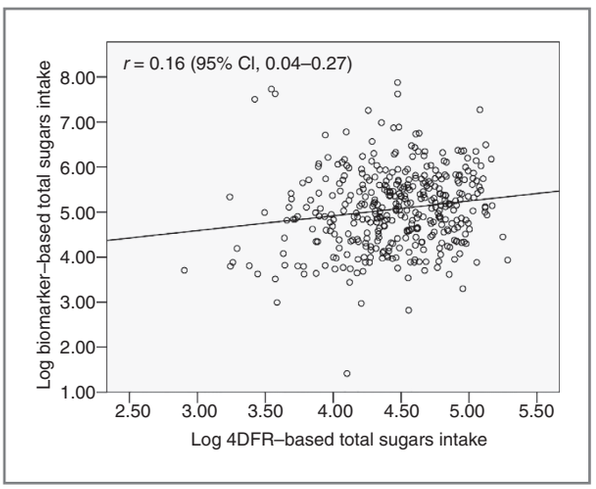

Figure 2. Association between log biomarker–based and log 4DFR–based total sugars intake among NPAAS participants.

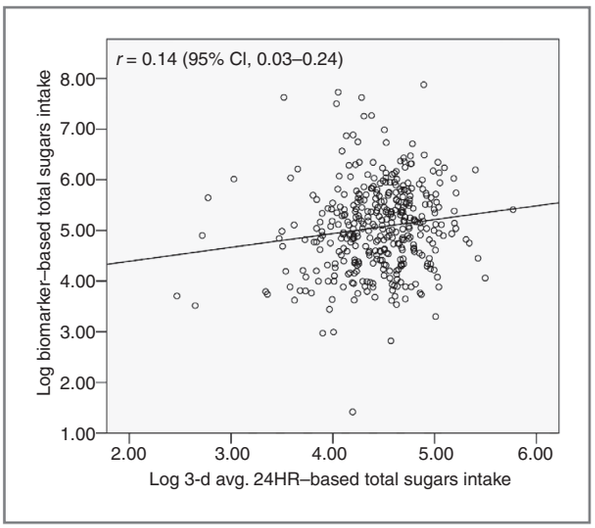

Figure 3. Association between log biomarker–based and log 3-d average 24HR–based total sugars intake among NPAAS participants.

attenuation, whereas AF < 1 indicates the extent of attenuation. Total sugars intake measured by one 24HR had the smallest attenuation factor (0.20), due to the large within-person errors in 24HR (Table 4); increasing the number of 24HRs increased the attenuation factor (two 24HR = 0.29; three 24HR = 0.34). The attenuation factor for 4DFR (0.33) was similar to the average of three 24HRs, whereas the attenuation factor for FFQ (0.22) was similar to one 24HR. Energy adjustment improved attenuation factors for FFQ (0.48) and three 24HR (0.57) but not 4DFR (0.32). Results for correlation coefficients were qualitatively similar to those observed for attenuation factors (Table 4).

Calibration equations to predict biomarker-based "true" sugars intake given self-reported intake and other covariates are presented in Table 5. For total sugars (g/d), self-reported intake alone explained little of the variation in "true" intake (adjusted $R^2 = 5\%–6\%$), and none of the other covariates was a strong predictor; among investigated covariates, only smoking status, education, and physical activity were retained (adjusted R^2 = 16%–18%). Using total sugars density modestly increased the percentage of variation explained by 24HR (18% vs. 6%) and FFQ (10% vs. 5%) but not 4DFR (7% vs. 6%). The final sugars density model explained 29% to 40% of variation in "true" intake, leaving a considerable portion unexplained. R^2 unadjusted for within-subject error in the biomarker had markedly lower values than adjusted R^2 across all instruments.

Excluding incomplete urines based on PABA had only modest effect on the biomarker-predicted intake estimates and measurement error parameters, attenuation factors, correlation coefficients, and regression coefficients for self-reported intakes (Supplementary Tables S1–S4). Omitting age from the measurement error model for the biomarker had a dramatic effect on the geometric means of the biomarker-predicted intake (making them considerably smaller and closer to the self-reported means; Supplementary Table S5), but little effect on other findings (Supplementary Tables S6–S8). Yet, geometric means measure intake on a *group level* only and give no indication of an instrument's ability to measure *individual's intake*, which is of interest in a cohort study.

Discussion

Using the urinary sugars biomarker to predict "true" intake of total sugars in these postmenopausal women, we found self-reported sugars to be substantially and roughly

Table 3. Measurement error structure for log total sugars intake and log total sugars density assessed by FFQ, 24HR, and 4DFR among NPAAS (2007–2009) participants (primary study, $n = 450$; reliability study, $n = 88$)

	Variance of true intake (σ_T^2)	Instrument	Slope in regression of reported on true intake (β_{Q1}, β_{F1}, or β_{R1})	Variance of person-specific bias ($\sigma_{u_Q}^2$, $\sigma_{u_F}^2$, or $\sigma_{u_R}^2$)	Correlation of person-specific bias with FFQ person-specific bias ($\rho_{u_Q}\rho_{u_F}$ or $\rho_{u_Q}\rho_{u_R}$)	Variance of within-person error ($\sigma_{\varepsilon_Q}^2$, $\sigma_{\varepsilon_F}^2$, or $\sigma_{\varepsilon_R}^2$)
Total sugars intake, g/d	0.26 (0.07)	FFQ	0.21 (0.09)[a]	0.15 (0.02)		0.06 (0.01)
		4DFR	0.20 (0.08)	0.07 (0.01)	0.62 (0.07)	0.06 (0.01)
		24HR	0.20 (0.08)	0.07 (0.01)	0.63 (0.07)	0.16 (0.01)
Total sugars density, g/1,000 kcal	0.23 (0.07)	FFQ	0.21 (0.08)	0.06 (0.01)		0.03 (0.01)
		4DFR	0.14 (0.07)	0.05 (0.01)	0.84 (0.07)	0.04 (0.01)
		24HR	0.22 (0.08)	0.04 (0.01)	0.70 (0.08)	0.11 (0.01)

NOTE: All the parameters were estimated using FFQ-, 24HR-, and 4DFR measurement errors models adjusted for BMI and age and biomarker measurement errors model adjusted for age.
[a]SE (all values in parentheses).

Table 4. Attenuation factors for self-reported intake and correlation coefficients between self-reported and true intakes for log total sugars and log total sugars density among NPAAS (2007–2009) participants (primary study, $n = 450$; reliability study, $n = 88$)

	Total sugars intake, g/d		Total sugars density, g/1,000 kcal	
Instrument	Attenuation factor	Correlation with true intake	Attenuation factor	Correlation with true intake
FFQ	0.22 (0.08)[a]	0.22 (0.08)	0.48 (0.13)	0.32 (0.09)
4DFR	0.33 (0.10)	0.26 (0.08)	0.32 (0.13)	0.21 (0.09)
Single 24HR	0.20 (0.06)	0.20 (0.06)	0.32 (0.07)	0.26 (0.07)
Avg. of two 24HR	0.29 (0.09)	0.24 (0.08)	0.48 (0.11)	0.32 (0.08)
Avg. of three 24HR	0.34 (0.10)	0.26 (0.08)	0.57 (0.13)	0.36 (0.09)

NOTE: All the parameters were estimated using FFQ-, 24HR-, and 4DFR measurement errors models adjusted for BMI and age and biomarker measurement errors model adjusted for age.

[a]SE (all values in parentheses).

equally misreported across the FFQ, 4DFR, and 24HR. This is the first study to assess measurement error properties and measurement errors–associated disease risk attenuation for food records with regard to sugars using a biomarker. Investigation of instruments' measurement error structure revealed considerable level of intake-related and person-specific biases, as well as within-subject error for all 3 instruments. Person-specific bias in 24HR and 4DFR was strongly correlated with bias in the FFQ. We found that sugars density measured by three 24HRs would result in the least attenuation of the association of sugars with disease risk, followed by FFQ and lastly by 4DFR. In this analysis, self-reported sugars and participants' personal characteristics recovered only some of the variation in "true" intake.

In this population, we observed large intake-related bias, comparable in magnitude across the 3 instruments. Interestingly, the bias did not improve with energy adjustment (even becoming greater) for the 4DFR, implying that other macronutrients may be larger contributors to intake-related bias in energy relative to sugars. The FFQ had almost double the person-specific bias compared with 4DFR and 24HR, which improved with energy adjustment. Another study of similar design using the sugars biomarker to investigate misreporting on FFQ and 24HR (16) reported similar levels of person-specific bias in females. Furthermore, similar to the OPEN study, we found the person-specific bias in 24HR strongly correlated with the bias in FFQ. The same was apparent for 4DFR, suggesting that the use of 24HR and 4DFR as reference instruments for FFQ-based total sugars may provide a very incomplete measurement error correction procedure. Whereas the intake-related bias, which was considerable in all 3 instruments, may ultimately cause inflation of the risk estimate through creating the flattened slope phenomenon (29), the person-specific bias and within-person random error have an opposing effect, causing risk attenuation (30). The fact that the AFs < 1 across all self-reports for both absolute and energy-adjusted sugars intake indicates that the person-specific bias and within-person random error overrode the effect of intake-related bias, causing underestimation of true effect in a disease model with self-reported sugars. The attenuation factor for three 24HRs was most favorable and much improved with energy adjustment (AF = 0.57), followed by the FFQ (AF = 0.48). In contrast, the attenuation factor for 4DFR became less favorable with energy adjustment (AF = 0.32), suggesting that errors in 4DFR-based sugars were independent from errors in the energy estimate. On the basis of the attenuation factors, in an association study using self-reports, a true RR = 2 for a given change in energy-adjusted intake will be observed as $2^{0.57} = 1.5$ for three 24HR, $2^{0.48} = 1.4$ for an FFQ, and $2^{0.32} = 1.2$ for 4DFR. In comparison to NPAAS women, OPEN women had lower attenuation factors (0.33 for FFQ and 0.35 for two 24HRs for sugars density).

Using calibration equations, self-report instruments explained very little variation in "true" sugars intake. BMI, age, and ethnicity have been identified as sources of systematic bias in dietary self-reporting, and their incorporation in equations may provide certain measurement errors adjustment and "strengthen the signal" (9, 10). None of the covariates we investigated was a particularly important predictor of "true" sugars intake. Our final models explained a total of 16% to 18% of "true" absolute sugars and 29% to 40% of "true" sugars density variability, lower than observed for energy, protein, and protein density in an earlier analysis of NPAAS (10). Being a smoker was associated with lower sugars consumption, whereas lower education with higher consumption, consistent with the latest reports on added sugars intake in the United States (31, 32). Although BMI would be expected to be a significant determinant of sugars intake and associated with sugars misreporting (33), we did not find BMI to improve predictability of the calibration equation. In the original (17) and a later feeding study (34), BMI showed no effect on the performance of the biomarker, that is, BMI was not a predictor (17) or an effect modifier (34) in the

Table 5. Regression calibration equations from regression of log-calibrated biomarker on log total sugars intake and log total sugars density assessed by FFQ, 4DFR, and three 24HRs among NPAAS (2007–2009) participants (primary study, $n = 384$; reliability study, $n = 78$)

	Total sugars intake, g/d			Total sugars density, g/1,000 kcal		
Variable	**FFQ**	**4DFR**	**Avg. of three 24HR**	**FFQ**	**4DFR**	**Avg. of three 24HR**
Intercept	3.51 (2.50)[a]	2.72 (2.50)	4.63 (2.45)	0.49 (2.53)	0.49 (2.51)	0.64 (2.49)
Log self-reported intake	0.23 (0.08)	0.30 (0.10)	0.26 (0.10)	0.44 (0.13)	0.36 (0.13)	0.52 (0.14)
Log age	0.17 (0.52)	0.25 (0.53)	−0.02 (0.52)	0.77 (0.53)	0.85 (0.54)	0.73 (0.53)
Log BMI	−0.11 (0.20)	−0.07 (0.20)	−0.19 (0.19)	−0.37 (0.20)	−0.38 (0.20)	−0.46 (0.20)
Smoking status (current)	−0.64 (0.24)	−0.63 (0.24)	−0.59 (0.24)	−0.56 (0.24)	−0.62 (0.25)	−0.57 (0.24)
Education						
<High school graduate	0.55 (0.23)	0.53 (0.23)	0.51 (0.22)	0.64 (0.24)	0.65 (0.24)	0.41 (0.24)
High school graduate	0.30 (0.14)	0.29 (0.14)	0.23 (0.14)	0.43 (0.15)	0.44 (0.15)	0.36 (0.14)
Some college	0.16 (0.09)	0.16 (0.09)	0.13 (0.09)	0.18 (0.09)	0.19 (0.09)	0.16 (0.09)
College graduate	0	0	0	0	0	0
Physical activity						
Square-root METs	0.04 (0.02)	0.04 (0.02)	—			
For self-reported intake only						
R^2	0.01	0.02	0.02	0.03	0.02	0.04
Adjusted[b] R^2	0.05	0.06	0.06	0.10	0.07	0.18
Multiple						
R^2	0.06	0.06	0.04	0.09	0.09	0.10
Adjusted[b] R^2	0.18	0.17	0.16	0.34	0.29	0.40

NOTE: Calibrated using the measurement error parameters generated from the feeding study (17). Regression calibration models automatically included self-reported intake, log age, and log BMI as covariates: additional covariates were chosen by backward selection ($P = 0.1$) from the following covariates: square root physical activity (METs/wk), smoking status (current, former/never), supplement use (yes, no), race (black, white, Hispanic, other), education (<high school graduate, high school graduate, some college, college graduate), annual income (<$20,000, $20,000–$49,999, ≥$50,000), and percentage of meals eaten at home (<75, 75–94, >94).
Abbreviation: METs, metabolic equivalents.
[a]SE (all values in parentheses).
[b]R^2 and multiple R^2 are adjusted for within-subject error in the biomarker.

association between the sugars biomarker and known intake. In contrast, a recent feeding study found gender and percentage body fat to be significant predictors in the regression of the sugars biomarker to known dietary sugars intake, in a study design originally developed to investigate the effect of high versus low glycemic load (GL) diets (35). Participants were fed 2 constant isocaloric diets, which were identical in macronutrient composition (% energy) and only differed by glycemic load and fiber content. Given all participants received the same diets adjusted to fit each participant's energy requirement, it may have been that gender and percentage body fat, being major determinants of energy requirement, were significant predictors of urinary sugars in this study (35).

Almost a third of our participants had urinary sucrose concentration below the LOD, similarly to what was found in the OPEN study (16). Sucrose is not a dominant sweetener in the U.S. diet, and it can be expected that urinary sucrose excretion will be low in this population.

Given the high burden of the study and the stratified recruitment algorithm that provided the intended enrollment for age, race, and BMI ranges, the response rate of 20.6% was judged to be adequate. Moreover, the intake estimates for energy, protein, and protein density, calculated by calibration equations generated in the Nutrient Biomarker Study (NBS), an earlier WHI biomarker study with a response rate almost double the rate in NPAAS, were highly correlated ($r = 0.95$–0.96), indicating good transferability of the calibration equations (9, 10). A potential limitation of this analysis is that our women were ≥60 years of age, whereas the study used to derive the sugars biomarker included subjects ages 23 to 66 years (17). Yet, excluding age from the calibrated biomarker did not appreciably change any of the findings. We acknowledge that the calibrated sugars biomarker was derived from a feeding study of a limited sample size (17). The biomarker contains certain level of bias, possibly arising from person-specific differences in absorption, gastrointestinal mucosal integrity, hepatic metabolism, or renal excretion of these 2 nutrients, determined by genetic factors, physiologic or medical conditions, and those may differ between different populations. Hence, more well-

designed feeding studies collecting multiple biomarker measurements per participants from different populations are needed to further explore the biomarker's characteristics and "behavior" and to confirm that the calibration parameters quantified from the original feeding study are stable and indeed applicable in different populations.

In summary, none of the self-report instruments provided a good estimate of sugars intake, although overall 24HRs seemed to perform the best. While 4DFR was comparable to multiple 24HRs in measuring absolute sugars, its performance did not improve with energy adjustment. Lower performance of the 4DFR compared with the 24HR and FFQ with regard to sugars density is surprising, given its administration most closely corresponded with the time frame of biomarker collection (i.e., within 2 weeks), whereas the 24HRs were administered a month apart starting 1 to 3 weeks after the biomarker assessment, and the FFQ, inquiring about usual diet over the preceding 3 months, was completed 2 weeks before biomarker collection.

Assuming that the calibrated sugars biomarker provides unbiased estimate of total sugars, this analysis suggests that measuring the biomarker in a subsample of the study population for calibration purposes may be necessary for obtaining unbiased risk estimates and reliable 95% CIs in disease association studies. Accordingly, the regression calibration equation for the FFQ developed here can be used in future WHI analyses to calibrate FFQ sugars intake and examine its associations with cancer and other disease outcomes.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Disclaimer

The study was approved by the institutional review boards of all participating institutions and has therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All participants gave their informed consent before their inclusion in the study.

Authors' Contributions

Conception and design: N. Tasevska, N. Potischman, M.L. Neuhouser, L. Van Horn, V. Kipnis

Development of methodology: N. Tasevska, D. Midthune, N. Potischman, J.W. Lampe, V. Kipnis

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): L.F. Tinker, J.W. Lampe, M.L. Neuhouser, L. Van Horn, R.L. Prentice

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): N. Tasevska, D. Midthune, N. Potischman, J.W. Lampe, R.L. Prentice, V. Kipnis

Writing, review, and/or revision of the manuscript: N. Tasevska, D. Midthune, L.F. Tinker, N. Potischman, J.W. Lampe, M.L. Neuhouser, J.M. Beasley, L. Van Horn, R.L. Prentice, V. Kipnis

Study supervision: N. Tasevska, L.F. Tinker, M.L. Neuhouser, L. Van Horn, V. Kipnis

Acknowledgments

The authors to thank the late Dr. Arthur Schatzkin for his support in pursuing this project and for providing the funding. A full listing of Women's Health Initiative investigators can be found at the following website: http://www.whi.org. A list of key investigators involved in this research follows. *Program Office:* (National Heart, Lung, and Blood Institute, Bethesda, MD) Jacques Rossouw, Shari Ludlam, Dale Burwen, Joan McGowan, Leslie Ford, and Nancy Geller. *Clinical Coordinating Center:* (Fred Hutchinson Cancer Research Center, Seattle, WA) Garnet Anderson, R.L. Prentice, Andrea LaCroix, and Charles Kooperberg. *Investigators and Academic Centers:* (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) JoAnn E. Manson; (MedStar Health Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Arizona, Tucson/Phoenix, AZ) Cynthia A. Thomson; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Iowa, Iowa City/Davenport, IA) Robert Wallace; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker. *Women's Health Initiative Memory Study:* (Wake Forest University School of Medicine) Sally Shumaker.

Grant Support

This work was supported by the Intramural Research Program of the National Cancer Institute, NIH, U.S. Department of Health and Human Services. The NPAAS was supported by the National Cancer Institute grant R01 CA119171-04A1 (L.F. Tinker, J.W. Lampe, M.L. Neuhouser, and R.L. Prentice). The WHI program is funded by the National Heart, Lung, and Blood Institute, NIH, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 28, 2014; revised August 21, 2014; accepted September 9, 2014; published OnlineFirst September 18, 2014.

References

1. WCRF/AICR. Food, nutrition, physical activity and the prevention of cancer: a global perspective. Washington, DC: AICR; 2007.
2. Tasevska N, Jiao L, Cross AJ, Kipnis V, Subar AF, Hollenbeck A, et al. Sugars in diet and risk of cancer in the NIH-AARP Diet and Health Study. Int J Cancer 2012;130:159–69.
3. Tasevska N, Park Y, Jiao L, Hollenbeck A, Subar AF, Potischman N. Sugars and risk of mortality in the NIH-AARP Diet and Health Study. Am J Clin Nutr 2014;99:1077–88.
4. Kipnis V, Freedman LS, Brown CC, Hartman AM, Schatzkin A, Wacholder S. Effect of measurement error on energy-adjustment models in nutritional epidemiology. Am J Epidemiol 1997;146:842–55.
5. Schatzkin A, Kipnis V. Could exposure assessment problems give us wrong answers to nutrition and cancer questions? J Natl Cancer Inst 2004;96:1564–5.
6. Kristal AR, Peters U, Potter JD. Is it time to abandon the food frequency questionnaire? Cancer Epidemiol Biomarkers Prev 2005;14:2826–8.
7. Schatzkin A, Kipnis V, Carroll RJ, Midthune D, Subar AF, Bingham S, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. Int J Epidemiol 2003;32:1054–62.
8. Day N, McKeown N, Wong M, Welch A, Bingham S. Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium. Int J Epidemiol 2001;30:309–17.
9. Neuhouser ML, Tinker L, Shaw PA, Schoeller D, Bingham SA, Horn LV, et al. Use of recovery biomarkers to calibrate nutrient consumption

self-reports in the Women's Health Initiative. Am J Epidemiol 2008;167:1247–59.
10. Prentice RL, Mossavar-Rahmani Y, Huang Y, Van Horn L, Beresford SA, Caan B, 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.
11. Jenab M, Slimani N, Bictash M, Ferrari P, Bingham SA. Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet 2009;125:507–25.
12. Kaaks R, Ferrari P. Dietary intake assessments in epidemiology: can we know what we are measuring? Ann Epidemiol 2006;16:377–80.
13. Prentice RL, Tinker LF, Huang Y, Neuhouser ML. Calibration of self-reported dietary measures using biomarkers: an approach to enhancing nutritional epidemiology reliability. Curr Atheroscler Rep 2013;15:353.
14. Kaaks R, Ferrari P, Ciampi A, Plummer M, Riboli E. Uses and limitations of statistical accounting for random error correlations, in the validation of dietary questionnaire assessments. Public Health Nutr 2002;5: 969–76.
15. Freedman LS, Midthune D, Carroll RJ, Tasevska N, Schatzkin A, Mares J, 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;174:1238–45.
16. Tasevska N, Midthune D, Potischman N, Subar AF, Cross AJ, Bingham SA, et al. Use of the predictive sugars biomarker to evaluate self-reported total sugars intake in the Observing Protein and Energy Nutrition (OPEN) study. Cancer Epidemiol Biomarkers Prev 2011;20: 490–500.
17. Tasevska N, Runswick SA, McTaggart A, Bingham SA. Urinary sucrose and fructose as biomarkers for sugar consumption. Cancer Epidemiol Biomarkers Prev 2005;14:1287–94.
18. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Control Clin Trials 1998;19:61–109.
19. Langer RD, White E, Lewis CE, Kotchen JM, Hendrix SL, Trevisan M. The Women's Health Initiative Observational Study: baseline characteristics of participants and reliability of baseline measures. Ann Epidemiol 2003;13:S107–21.
20. Johnson-Kozlow M, Rock CL, Gilpin EA, Hollenbach KA, Pierce JP. Validation of the WHI brief physical activity questionnaire among women diagnosed with breast cancer. Am J Health Behav 2007;31:193–202.
21. Patterson RE, Kristal AR, Tinker LF, Carter RA, Bolton MP, Agurs-Collins T. Measurement characteristics of the Women's Health Initiative food frequency questionnaire. Ann Epidemiol 1999;9:178–87.
22. 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. Am J Clin Nutr 2008;88:324–32.
23. Bingham S, Cummings JH. The use of 4-aminobenzoic acid as a marker to validate the completeness of 24 h urine collections in man. Clin Sci (Lond) 1983;64:629–35.
24. 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–54S.
25. Schoeller DA. Recent advances from application of doubly labeled water to measurement of human energy expenditure. J Nutr 1999;129:1765–8.
26. Willett W. Commentary: Dietary diaries versus food frequency questionnaires-a case of undigestible data. Int J Epidemiol 2001;30:317–9.
27. Mossavar-Rahmani Y, Tinker LF, Huang Y, Neuhouser ML, McCann SE, Seguin RA, et al. Factors relating to eating style, social desirability, body image and eating meals at home increase the precision of calibration equations correcting self-report measures of diet using recovery biomarkers: findings from the Women's Health Initiative. Nutr J 2013;12:63.
28. Neuhouser ML, Di C, Tinker LF, Thomson C, Sternfeld B, Mossavar-Rahmani Y, et al. Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. Am J Epidemiol 2013;177:576–85.
29. Wacholder S. When measurement errors correlate with truth: surprising effects of nondifferential misclassification. Epidemiology 1995;6:157–61.
30. Thiebaut AC, Freedman LS, Carroll RJ, Kipnis V. Is it necessary to correct for measurement error in nutritional epidemiology? Ann Intern Med 2007;146:65–7.
31. Ervin RB, Ogden CL. Consumption of added sugars among U.S. adults, 2005–2010. Hyattsville, MD: National Center for Health Statistics; 2013.
32. Thompson FE, McNeel TS, Dowling EC, Midthune D, Morrissette M, Zeruto CA. Interrelationships of added sugars intake, socioeconomic status, and race/ethnicity in adults in the United States: National Health Interview Survey, 2005. J Am Diet Assoc 2009;109:1376–83.
33. Bingham S, Luben R, Welch A, Tasevska N, Wareham N, Khaw KT. Epidemiologic assessment of sugars consumption using biomarkers: comparisons of obese and nonobese individuals in the European prospective investigation of cancer Norfolk. Cancer Epidemiol Biomarkers Prev 2007;16:1651–4.
34. Joosen AM, Kuhnle GG, Runswick SA, Bingham SA. Urinary sucrose and fructose as biomarkers of sugar consumption: comparison of normal weight and obese volunteers. Int J Obes (Lond) 2008;32: 1736–40.
35. Song X, Navarro SL, Diep P, Thomas WK, Razmpoosh EC, Schwarz Y, et al. Comparison and validation of 2 analytical methods for measurement of urinary sucrose and fructose excretion. Nutr Res 2013;33:696–703.