Practice of Epidemiology

Relative Validity of Nutrient Intakes Assessed by Questionnaire, 24-Hour Recalls, and Diet Records as Compared With Urinary Recovery and Plasma Concentration Biomarkers: Findings for Women

Changzheng Yuan, Donna Spiegelman, Eric B. Rimm, Bernard A. Rosner, Meir J. Stampfer, Junaidah B. Barnett, Jorge E. Chavarro, Jennifer C. Rood, Lisa J. Harnack, Laura K. Sampson, and Walter C. Willett*

* Correspondence to Dr. Walter C. Willett, Department of Nutrition, Harvard T.H. Chan School of Public Health, Building II, 3rd Floor, 655 Huntington Avenue, Boston, MA 02115 (e-mail: wwillett@hsph.harvard.edu).

We evaluated the performance of a semiquantitative food frequency questionnaire (SFFQ), the Automated Self-Administered 24-Hour Dietary Recall (ASA24), and 7-day dietary records (7DDRs), in comparison with biomarkers, in the estimation of nutrient intakes among 627 women in the Women's Lifestyle Validation Study (United States, 2010–2012). Two paper SFFQs, 1 Web-based SFFQ, 4 ASA24s (beta version), 2 7DDRs, 4 24-hour urine samples, 1 doubly labeled water measurement (repeated among 76 participants), and 2 fasting blood samples were collected over a 15-month period. The dietary variables evaluated were energy, energy-adjusted intakes of protein, sodium, potassium, and specific fatty acids, carotenoids, α-tocopherol, retinol, and folate. In general, relative to biomarkers, averaged ASA24s had lower validity than the SFFQ completed at the end of the data-collection year (SFFQ2); SFFQ2 had slightly lower validity than 1 7DDR; the averaged SFFQs had validity similar to that of 1 7DDR; and the averaged 7DDRs had the highest validity. The deattenuated correlation of energy-adjusted protein intake assessed by SFFQ2 with its biomarker was 0.46, similar to its correlation with 7DDRs (deattenuated $r = 0.54$). These data indicate that the SFFQ2 provides reasonably valid measurements of energy-adjusted intake for most of the nutrients assessed in our study, consistent with earlier conclusions derived using 7DDRs as the comparison method. The ASA24 needs further evaluation for use in large population studies, but an average of 3 days of measurement will not be sufficient for some important nutrients.

biomarkers; concentration biomarkers; diet records; food frequency questionnaires; nutrient intakes; recall; recovery biomarkers; relative validity

In large cohort studies of diet and health outcomes, food frequency questionnaires (FFQs) have been the primary method of dietary assessment because they can be self-administered, can be efficiently processed, and provide data on individual intakes of both foods and nutrients over an extended period of time (1). The validity of these questionnaires for measuring long-term diet has been evaluated extensively in many studies, indicating a moderate-to-strong concordance with dietary records or 24-hour dietary recalls (1, 2). However, correlated errors between FFQs, diet records, and recalls may tend to overestimate

the validity of FFQs. Conversely, both comparison methods have their own sources of error, which would tend to underestimate the validity of the FFQ. To overcome these limitations, biomarkers have been utilized as the reference method for evaluating the performance of other dietary methods (2).

Recovery biomarkers of absolute intake, such as urinary nitrogen (to assess protein intake), potassium, and sodium levels, are based on the balance between intake and output; they estimate absolute intakes over a certain time period (3). Concentration biomarkers reflect dietary composition but

cannot be directly translated to absolute intakes because blood levels are influenced by numerous physiological and environmental factors in addition to diet (3, 4). Despite these limitations, concentration biomarkers can serve as objective reference measures for assessment of the relative validity of different dietary assessment methods (3, 5–7).

Many investigators have evaluated the performance of FFQs and 24-hour recalls against intake biomarkers. Most recently, in a pooled analysis of 5 validation studies, Freedman et al. (8, 9) examined correlations between FFQ or 24-hour recalls and biomarkers for energy, protein, sodium, and potassium. In general, FFQs had stronger correlations with biomarkers for protein density (percentage of energy derived from protein) than for absolute protein intake; multiple 24-hour recalls performed better than a single recall in measuring energy and protein density; and modest correlations were observed between energy-adjusted sodium and potassium intakes assessed by FFQs or multiple 24-hour recalls and biomarkers (6, 8–17). Studies that evaluated the performance of a semiquantitative FFQ (SFFQ) against concentration biomarkers documented that intakes of specific carotenoids, fatty acids, and vitamins were poorly-to-moderately correlated with their corresponding biological fluid or tissue levels (18–25). Interpretation of findings from validation studies using biomarkers requires careful consideration of their variation in diet by season, the time frame of data collection, and the temporal relationships among the dietary methods (3, 6, 26–28). Limitations in most existing studies include the facts that few nutrient biomarkers were evaluated and that replicate biomarker measurements were taken only in small subgroups, if at all, with a relatively short time interval between measurements, which would not adequately account for within-person variation over a year (29–31). Furthermore, the biomarker measurements were often collected close in time to 24-hour recalls, which may have produced overestimation of the validity of the 24-hour recalls due to correlated within-person variation between the 2 methods, if they both reflect short-term intake.

To address these issues, we evaluated the relative validity of multiple self-reporting dietary assessment methods over a 15-month period in comparison with repeated biomarker measurements. Different dietary methods were used at least several weeks apart and in random order to avoid artificially high correlations. We evaluated the performance of paper and Web versions of an SFFQ, the Harvard or Willett FFQ (32–34), which has been widely used in many cohort studies; the Automated Self-Administered 24-Hour Dietary Recall (ASA24; beta version) developed by the National Cancer Institute (35); and 7-day dietary records (7DDRs), a commonly used "gold standard" method. The study participants completed 2 paper SFFQs (SFFQ1, which was completed at study baseline, and SFFQ2, completed at the end of the data-collection year) and 1 online SFFQ 2 weeks before or after completion of the second SFFQ.

METHODS

The current analysis was based on the Women's Lifestyle Validation Study, one of 3 studies that comprise the Multi-Cohort Eating and Activity Study for Understanding Reporting Error (MEASURE), designed to investigate the measurement error structure associated with self-reported dietary and physical activity assessments (36). The Women's Lifestyle Validation Study was conducted within the Nurses' Health Study (37) and Nurses' Health Study II (38, 39). The study has been described in detail previously (40). This analysis was approved by the human subjects committees of the Harvard T.H. Chan School of Public Health and Brigham and Women's Hospital.

To represent the 1-year period typically used as the time frame for dietary questionnaires, we spread the dietary and biomarker measurements out over a period of approximately 15 months, altered the order of measurements into 4 groups, and randomized participants into these groups (Figure 1). The 15-month study period was divided into 5 phases, with each phase representing a 3-month interval. In addition to the self-reported
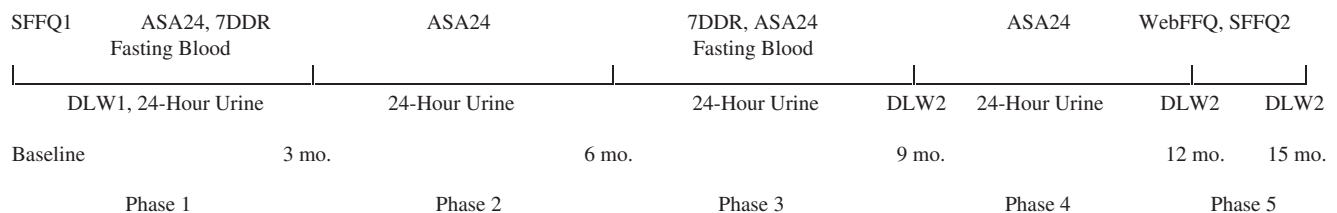
Figure 1. Timeline of the dietary assessment activities undertaken by group 1 participants in the Women's Lifestyle Validation Study, United States, 2010–2012. The study participants completed 2 paper semiquantitative food frequency questionnaires (SFFQs)—SFFQ1, which was completed at study baseline, and SFFQ2, completed at the end of the data-collection year—and 1 Web-based food frequency questionnaire (WebFFQ) 2 weeks before or after completion of the second SFFQ. Group 3 had the same data collection timeline as group 1; groups 2 and 4 were assigned similar data collection timelines as group 1, except that groups 2 and 4 were asked to complete 7-day dietary records (7DDRs) and give fasting blood samples in study phases 2 and 4 instead. The first doubly labeled water (DLW) measurement (DLW1) was completed in phases 1, 2, 3, and 4 of the study by women in groups 1, 2, 3, and 4, respectively. A subgroup of women in group 1 completed a second DLW measurement (DLW2) at 9, 12, or 15 months (mo.). Within the same phase, 7DDRs, Automated Self-Administered 24-Hour Dietary Recalls (ASA24s), DLW, 24-hour urine samples, and blood samples were completed 1–5 weeks apart to avoid artificially high correlations. For groups 1 and 3, the ASA24 was completed first, followed by the 7DDR in phase 1, but this order was reversed in phase 3. For groups 2 and 4, the 7DDR was completed first, followed by the ASA24 in phase 2, but this order was reversed in phase 4. Additionally, groups 1 and 3 completed the WebFFQ about 2 weeks before completion of SFFQ2, and groups 2 and 4 about 2 weeks after completion of SFFQ2. To minimize alteration in eating behavior, the participants were not told in advance the day on which they would be asked to complete the ASA24; days were randomly selected and may or may not have included a weekend day.

dietary data, participants also underwent doubly labeled water (DLW) assessment (once for all participants randomly over a period of 1 year, with a subset of women having the first measurement in phase 1 and a repeated measurement 6–12 months later), gave 4 24-hour urine samples (once each season), and had 2 blood samples taken approximately 6 months apart. By design, within the same phase of the study, the 7DDRs, ASA24s, DLW, 24-hour urine samples, and blood samples were collected 1–5 weeks apart from each other and in random order to avoid artificially high correlations (6, 41). Notably, not all participants had the designed number of measurements of each method due to both lower participation rates for the ASA24 and budget constraints for analysis of samples. Among 796 participants enrolled, we excluded those with SFFQ energy intakes less than 600 kcal/day or more than 3,500 kcal/day or with more than 70 blank SFFQ items (24 for SFFQ1 and 17 for SFFQ2). Our primary analysis included 627 participants with complete data for SFFQ1, SFFQ2, a Web-based version of the paper SFFQ that was enhanced by the use of branched questions (WebFFQ), and at least 1 of the following: 7DDR, ASA24, DLW, 24-hour urinary measurement, and fasting blood measurement.

Self-reported dietary methods

As described previously, the 152-food-item paper SFFQs were mailed to and completed by the participants. 7DDRs were completed by the participants, each of whom received an Escali food scale (Escali Corporation, Burnsville, Minnesota) and ruler, an instructional DVD, and instructions via telephone for keeping the 7DDRs. Participants measured and reported gram weights for foods before and after eating so actual intake could be computed, and they provided recipes of all home-prepared foods, including the number of servings in each recipe and the portion of the recipe they had consumed. Dietary supplements taken were also recorded. Participants self-administered the WebFFQ and ASA24 online. Because the beta version of ASA24 did not include supplement intake, we could only use ASA24 nutrients from food in our analyses. The ASA24 data were cleaned by the National Cancer Institute group that developed the ASA24, using their standard process to detect errors and extreme values.

For the SFFQ and WebFFQ, we derived values for fatty acid variables using the updated nutrient compositions of commonly consumed brands and types of margarines, cooking fats, and processed foods analyzed in the same laboratory as that used for our plasma fatty acid analyses; this updated database (the Harvard Food Composition Table) was used in this study. 7DDR fatty acid variables were derived based on the Nutrition Data System for Research 2011 database (42, 43); therefore, data for some specific isomers of certain fatty acids were not derived. Results for fatty acid composition focused on total saturated fatty acids, total monounsaturated fatty acids, total polyunsaturated fatty acids, total *trans*-fatty acids, and several specific polyunsaturated fatty acids (linoleic acid, α-linolenic acid, and the long-chain n-3 fatty acids). Carotenoids and tocopherols analyzed in this study included α-carotene, β-carotene, β-cryptoxanthin, lutein-zeaxanthin, lycopene, α-tocopherol, and γ-tocopherol; we also included retinol. The dietary folate equivalent value (the sum of natural

food folate (μg) and 1.7 × total synthetic folic acid (μg)) was compared with the plasma folate level. Nutrients assessed by means of the ASA24 did not account for supplements; however, the amounts used for most types of fatty acids (except n-3 fatty acids), α-carotene, β-cryptoxanthin, lutein-zeaxanthin, and lycopene were quite small and could be considered ignorable in comparison with intake from food. For long-chain n-3 fatty acids, β-carotene, retinol, α-tocopherol, and folate, we conducted separate analyses for women not taking a supplement containing the corresponding nutrient.

Biomarker measurements

DLW is considered the gold standard for measurement of total energy expenditure among free-living persons. Total daily energy expenditure (kcal/day) was obtained from DLW data. Biochemical indicators of protein (g/day), sodium (mg/day), and potassium (mg/day) intakes were obtained from 24-hour urinary data. Urinary values for nitrogen, sodium, and potassium were calculated by multiplying urinary concentration by reported total urine volume. Urinary nitrogen level (in grams) was divided by 0.81 to convert the measurement to dietary nitrogen and was then multiplied by 6.25 to obtain dietary protein intake (g/day) (44). Urinary values were divided by 0.86 for sodium and 0.80 for potassium to convert them to the corresponding dietary values (mg/day) (9, 45). We collected plasma samples to measure specific fatty acids and carotenoids, retinol, tocopherols, and folate. Standard measures of blood lipids (total cholesterol, high-density lipoprotein cholesterol, and triglycerides) were used to adjust for variation in carotenoids, retinol, and tocopherols. Methods used for biomarker assessment are described in detail in the Web Appendix (available at https://academic.oup.com/aje).

In the final analysis, we included 627 participants with data from SFFQ2, SFFQ1, the WebFFQ, and at least 1 of the following methods: 7DDR (626 women completed weeks 1 and 2), ASA24 (80, 117, 196, and 234 women completed 1, 2, 3, and 4 ASA24s, respectively), 24-hour urinary measurement (607, 600, 605, and 609 women completed days 1, 2, 3, and 4, respectively), fasting blood measurement (for fatty acids/carotenoids, 615 and 306 women completed blood samples 1 and 2, respectively; for folate, 456 completed blood samples 1 and 2), or DLW measurement (624 women with measurements spread out over the year; 76 underwent a second DLW measurement approximately 6–12 months later).

Statistical analysis

Average daily intakes of energy, protein, sodium, potassium, specific fatty acids, carotenoids, tocopherols, and folate estimated from the SFFQs, ASA24s (mean = 2.9 days), and 7DDRs (mean = 14 days) were compared with converted absolute values or concentrations measured by the corresponding biomarkers. Log transformation was performed for nutrients and biomarkers to increase normality. Sodium:potassium ratio was calculated as the sodium value divided by the potassium value. Energy-adjusted intakes are of greatest importance because people alter their intakes of specific nutrients primarily by changing the composition of their diet, keeping total energy constant (1, 46). Therefore, for recovery biomarkers and associated self-reported nutrients, we

Table 1.   Characteristics of Participants in the Women's Lifestyle Validation Study (Data Provided by 627 US Female Nurses Aged 45–80 Years), 2010–2012

| Variable[a] | Overall[b] (n = 627) | Energy[c] | | Fatty Acids and Carotenoids | | Folate (2 Blood Samples) (n = 456) |
		1 DLW (n = 548)	2 DLWs (n = 76)	1 Blood Sample (n = 321)	2 Blood Samples (n = 306)	
Age at enrollment, years	61.4 (9.5)	61.5 (9.5)	60.8 (9.5)	62.3 (9.9)	60.5 (9.1)	60.8 (9.3)
Height, m	1.64 (0.07)	1.64 (0.07)	1.63 (0.07)	1.64 (0.07)	1.64 (0.07)	1.64 (0.07)
Weight, kg	71.5 (15.5)	71.9 (15.6)	69.7 (14.6)	71.6 (15.6)	72.7 (15.5)	71.5 (15.3)
Weight change[d], kg	−0.2 (2.6)	−0.2 (2.6)	−0.4 (3.3)	−0.1 (2.6)	−0.3 (2.6)	−0.2 (2.7)
Body mass index[e]	26.5 (5.4)	26.6 (5.5)	26.2 (5.0)	26.6 (5.4)	26.9 (5.4)	26.5 (5.4)
White race, %	90.6	89.2	97.4	92.8	88.2	91.0
Current smoking, %	1.9	1.5	5.3	1.0	3.0	2.2

Abbreviation: DLW, doubly labeled water.

[a] Values are presented as mean (standard deviation) unless otherwise indicated.

[b] Women's Lifestyle Validation Study participants included 270 participants from the Nurses' Health Study (with an average age of 70.9 years; 1.1% current smokers) and 357 participants from Nurses' Health Study II (with an average age of 54.3 years; 2.5% current smokers).

[c] The first DLW was completed over a 1-year period ($n = 164$, $n = 162$, $n = 153$, and $n = 145$ in each phase of the study; total $n = 624$). The second DLW was completed at 9 months ($n = 28$), 12 months ($n = 26$), or 15 months ($n = 22$).

[d] Weight change was estimated over the 12 months of the study period.

[e] Weight (kg)/height (m)^2.

calculated energy-adjusted protein intake (percentage of total energy derived from protein) and sodium and potassium intakes (the ratio of nutrient intake (mg) to energy intake (1,000 kcal)) (1). Notably, biomarker energy-adjusted nutrient values were calculated using DLW measurements collected during the same study phase with urinary recovery biomarkers, which led to a reduced sample size in each phase and the use of 1 24-hour urine measurement for most participants.

For concentration biomarkers, to remove variation in plasma nutrient levels due to nondietary factors, we obtained fatty acid residuals from multivariate linear regression of each fatty acid composition on the following covariates: age (years), body mass index (weight (kg)/height (m)^2) at enrollment, current weight (pounds), current smoking status (yes, no), menopausal status (yes, no), use of hormone replacement within the past 6 months (yes, no), and fasting status (≥8 hours, <8 hours) at each blood drawing; fat-soluble carotenoid, retinol, and tocopherol residuals were calculated by additionally adjusting for levels of plasma lipids (total triglycerides and total cholesterol). For self-reported nutrient intakes that were compared with concentration biomarkers, residuals were calculated using models adjusting for total energy intake, age, body mass index at enrollment, current weight, and current smoking status. To further reduce the influence of extreme nutrient distributions, the subsequent analyses were based on the rank scale.

To assess the reproducibility of repeated SFFQ, ASA24, 7DDR, and biomarker measurements, we calculated rank intraclass correlation coefficients (ICCs) for each nutrient. For assessment of the relative validity of each method, the average of the biomarker measurements, repeated at an interval of at least several months over a 15-month period, served as the objective comparison method. We calculated Spearman correlation coefficients (r_s) and their 95% confidence intervals for correlations between nutrient intakes reported on single and averaged ASA24s, SFFQs, and 7DDRs and the corresponding intakes

measured by biomarkers. In general, correlations of 0.4–0.6 (moderate correlations) with a relative "gold standard" are considered to represent a reasonably valid measurement (6). Because random within-person variation in the biomarker measurements will attenuate these correlations, we deattenuated correlation coefficients to reduce the effect of random error in the comparison methods (47, 48), using a method to account for the variable number of repeats of the comparison method (48–51). Linear mixed models that account for dependent correlation coefficients were used to evaluate the relative validity of nutrients assessed by means of different self-reporting methods.

RESULTS

At baseline, participants had a mean age of 61 years and a mean body mass index of 26.5. Participants were predominantly white (91%), and 2% were current smokers. The subgroups with different repeated measurements of biomarkers had characteristics similar to those of the overall population (Table 1).

The distributions of mean daily nutrient intakes or concentrations are shown in Table 2. Compared with biomarkers, total energy intake, sodium intake, and sodium density were underreported by all self-reporting methods. Notably, the SFFQ was not targeted on sodium and did not seek information on added salt. SFFQ-assessed protein intake was similar to that of the biomarkers; however, 7DDR and ASA24 tended to underestimate protein intake. Protein density was slightly overestimated by these methods due to greater underreporting of total energy intake than of protein intake. Potassium intake was underreported by both ASA24 and 7DDR and was slightly underreported by SFFQ. The ASA24-assessed sodium:potassium ratio was similar to that of the biomarkers; however, the SFFQ and 7DDR tended to underreport sodium:potassium ratio. Fatty

Table 2. Mean (Standard Deviation) Nutrient Intakes Estimated by Means of Biomarker Measurements and 3 Self-Reporting Methods (SFFQ2, 7DDRs, and ASA24s), Women's Lifestyle Validation Study, 2010–2012

Nutrient[a]	No. of Women	Dietary Assessment Method			
		Biomarkers	SFFQ2[b]	7DDRs	ASA24s
Total energy, kcal/day	624	2,195 (369)	1,853 (523)	1,737 (335)	1,815 (471)
Protein, g/day	624	79.8 (18.9)	80.4 (23.9)	72.4 (16.1)	77.2 (23.7)
Protein density[c], % of energy	624	14.9 (4.0)	17.5 (3.0)	16.9 (2.9)	17.4 (3.8)
Sodium, mg/day	624	3,388 (1,004)	2,060 (661)	2,645 (641)	3,079 (904)
Sodium density[c], mg/1,000 kcal	624	1,590 (581)	1,121 (215)	1,536 (277)	1,742 (409)
Potassium, mg/day	624	3,043 (866)	3,286 (966)	2,661 (655)	2,788 (856)
Potassium density[c], mg/1,000 kcal	624	1,433 (515)	1,794 (290)	1,551 (316)	1,585 (376)
Sodium:potassium ratio	624	1.2 (0.4)	0.6 (0.2)	1.0 (0.3)	1.2 (0.4)
Fatty acids, % of total[d]					
Saturated fatty acids	627	31.4 (2.7)	35.6 (6.2)	33.1 (4.7)	33.8 (6.2)
Monounsaturated fatty acids	627	22.2 (2.8)	38.0 (4.2)	36.1 (3.1)	35.6 (3.8)
PUFAs					
α-Linolenic acid (18:3n-3c)	627	0.6 (0.2)	2.3 (1.1)	2.3 (1.0)	2.2 (1.0)
Long-chain n-3 fatty acids (DHA + DPA + EPA)	627	3.5 (1.6)	0.8 (0.7)	0.5 (0.7)	N/A
Long-chain n-3 fatty acid NS[e]	363	2.8 (1.0)	0.4 (0.4)	0.3 (0.3)	0.3 (0.6)
Linoleic acid (18:2n-6cc)	627	30.0 (3.7)	18.1 (3.7)	19.5 (3.7)	19.0 (4.9)
Arachidonic acid (20:4n-6c)	627	8.2 (1.9))	0.2 (0.1)	0.2 (0.1)	0.2 (0.1)
Total PUFAs	627	44.8 (4.0)	22.2 (4.6)	22.4 (4.2)	21.9 (5.6)
Trans-fatty acids	627	1.1 (0.4)	2.7 (0.7)	3.3 (1.1)	N/A
Carotenoids, μg/L or μg/day[f]					
α-Carotene	627	101 (94)	855 (767)	613 (490)	498 (771)
β-Carotene	627	397 (398)	6,498 (4,055)	4,291 (2,828)	N/A
β-Carotene NS[e]	335	348 (349)	6,342 (3,844)	3,987 (2,707)	3,434 (2,966)
β-Cryptoxanthin	627	160 (142)	111 (96)	154 (182)	100 (138)
Lutein-zeaxanthin	627	284 (146)	3,926 (3,182)	2,700 (2,799)	2,598 (2,940)
Lycopene	627	431 (172)	5,613 (3,987)	4,986 (3,570)	5,091 (5,412)
Retinol, μg/L or μg/day[f]					
Retinol activity equivalents	627	620 (136)	1,810 (1,302)	1,480 (1,297)	N/A
Retinol activity equivalent NS[e]	207	591 (126)	931 (383)	787 (399)	762 (605)
Tocopherols, mg/L or mg/day[g]					
α-Tocopherol	627	14.1 (4.8)	42.3 (71.9)	37.3 (54.1)	N/A
α-Tocopherol NS[e]	148	11.4 (2.7)	9.3 (4.7)	9.0 (3.6)	7.8 (3.8)
γ-Tocopherol	627	1.4 (0.8)	10.1 (4.3)	11.2 (3.9)	N/A
Folate, ng/mL or μg/day[h]					
Dietary folate equivalents	456	31.9 (21.6)	1,150 (664)	1,073 (629	N/A
Dietary folate equivalent NS[e]	134	19.3 (16.1)	558 (375)	518 (216)	487(212)

Abbreviations: ASA24, Automated-Self-Administered 24-Hour Dietary Recall; 7DDR, 7-day dietary record; DHA, docosahexaenoic acid; DPA, docosapentaenoic acid; EPA, eicosapentaenoic acid; N/A, not applicable; NS, nonsupplement subgroup; PUFAs, polyunsaturated fatty acids; SD, standard deviation; SFFQ, semiquantitative food frequency questionnaire.

[a] Nutrient values are presented using the average of repeated measures for each method: 627 participants completed an average of 1.1 doubly labeled water measurements, 3.9 urinary biomarker measurements, 1.5 measurements of all plasma biomarkers except folate, 1 SFFQ2, 2 weeks of 7DDR, and an average of 2.9 days of ASA24. A total of 456 participants completed 2 plasma folate biomarker measurements. Mean values and standard deviations were calculated using untransformed variables.

[b] SFFQ2 was the SFFQ completed at the end of the data-collection year.

[c] Nutrient density was calculated using the concurrent energy assessment.

[d] For concentration biomarkers, the nutrient units for biomarkers and self-reporting methods are not directly comparable. The unit for fatty acid biomarkers is percentage of total fatty acids, while the unit for self-reported fatty acid intake is percentage of total fat.

[e] Subgroup of women not taking a supplement containing the corresponding nutrient.

[f] The unit for carotenoid/retinol biomarkers is μg/L, while the unit for self-reported carotenoid/retinol intake is μg/day.

[g] The unit for tocopherol biomarkers is mg/L, while the unit for self-reported tocopherol intake is mg/day.

[h] The unit for folate biomarkers is ng/mL, while the unit for self-reported folate intake is μg/day.

Table 3. Reproducibility (Rank Intraclass Correlation) of Nutrient Intakes Estimated by Means of Biomarker Measurements and 3 Self-Reporting Methods (SFFQs, 7DDRs, and ASA24s), Women's Lifestyle Validation Study, 2010–2012

Nutrient	No. of Women	Dietary Assessment Method[a]							
		Biomarkers		SFFQs		7DDRs		ASA24s	
		Unadjusted	Adjusted[b]	Unadjusted	Adjusted[b]	Unadjusted	Adjusted[b]	Unadjusted	Adjusted[b]
Total energy, kcal/day	624	0.71	N/A	0.70	N/A	0.63	N/A	0.28	N/A
Protein, g/day	624	0.56	0.33	0.65	0.63	0.64	0.60	0.25	0.21
Sodium, mg/day	624	0.37	0.26	0.73	0.68	0.58	0.50	0.23	0.20
Potassium, mg/day	624	0.49	0.56	0.71	0.69	0.75	0.70	0.37	0.32
Fatty acids, % of total fatty acids									
Saturated fatty acids	627	0.21	0.21	0.65	0.64	0.58	0.58	0.19	0.19
MUFAs	627	0.69	0.62	0.62	0.62	0.41	0.41	0.11	0.11
PUFAs									
α-Linolenic acid (18:3n-3c)	627	0.59	0.58	0.59	0.58	0.48	0.47	0.18	0.17
Long-chain n-3 fatty acids (DHA + DPA + EPA)	627	0.78	0.74	0.75	0.73	0.61	0.59	N/A	N/A
Long-chain n-3 fatty acid NS[c]	363	0.67	0.63	0.69	0.66	0.53	0.52	0.11	0.11
Linoleic acid (18:2n-6cc)	627	0.58	0.53	0.54	0.53	0.44	0.44	0.14	0.13
Total PUFAs	627	0.43	0.38	0.58	0.58	0.46	0.46	0.15	0.14
Trans-fatty acids	627	0.25	0.23	0.65	0.65	0.41	0.40	N/A	N/A
Carotenoids, µg/L or µg/day[d]									
α-Carotene	627	0.83	0.78	0.69	0.65	0.37	0.36	0.15	0.12
β-Carotene	627	0.86	0.81	0.70	0.66	0.47	0.44	N/A	N/A
β-Carotene NS[c]	335	0.85	0.79	0.69	0.68	0.52	0.49	0.25	0.16
Lutein-zeaxanthin	627	0.81	0.74	0.71	0.69	0.53	0.50	0.24	0.20
β-Cryptoxanthin	627	0.77	0.71	0.74	0.71	0.32	0.29	0.19	0.14
Lycopene	627	0.65	0.56	0.63	0.55	0.24	0.22	0.11	0.11
Retinol, µg/L or µg/day[d]									
Retinol activity equivalents	627	0.68	0.64	0.68	0.68	0.65	0.64	N/A	N/A
Retinol activity equivalent NS[c]	207	0.63	0.58	0.54	0.56	0.48	0.45	0.21	0.14
Tocopherols, mg/L or mg/day[e]									
α-Tocopherol	627	0.63	0.52	0.71	0.72	0.78	0.77	N/A	N/A
α-Tocopherol NS[c]	148	0.56	0.43	0.47	0.50	0.58	0.65	0.24	0.20
γ-Tocopherol	627	0.71	0.67	0.66	0.60	0.44	0.38	N/A	N/A
Folate, ng/mL or µg/day[f]									
Dietary folate equivalents	456	0.70	0.68	0.66	0.64	0.78	0.73	N/A	N/A
Dietary folate equivalent NS[c]	134	0.73	0.73	0.40	0.37	0.69	0.68	0.21	0.18

Abbreviations: ASA24, Automated Self-Administered 24-Hour Dietary Recall; 7DDR, 7-day dietary record; DHA, docosahexaenoic acid; DPA, docosapentaenoic acid; EPA, eicosapentaenoic acid; MUFAs, monounsaturated fatty acids; N/A, not applicable; NS, nonsupplement subgroup; PUFAs, polyunsaturated fatty acids; SFFQ, semiquantitative food frequency questionnaire.

[a] Two energy biomarkers (doubly labeled water), 6–12 months apart; 4 urinary biomarker measurements, taken approximately every 3 months over 1 year; 2 plasma biomarker measurements, taken approximately 6 months apart; 2 SFFQs, administered approximately 1 year apart; 2 7DDRs, completed approximately 6 months apart; 4 ASA24s, completed approximately every 3 months over 1 year.

[b] Protein, sodium, and potassium values from self-reporting measures and recovery biomarkers were adjusted for total energy intake using the energy density method. Values for concentration biomarkers and associated self-reported nutrients were adjusted for age and body mass index at enrollment, current weight, and smoking status at each measurement. Self-reported nutrient values were further adjusted for total energy intake (except for fatty acid composition). Concentration biomarker values were further adjusted for postmenopausal status, hormone use, and fasting status at blood drawing. Plasma carotenoid, retinol, and tocopherol values were additionally adjusted for plasma lipid levels.

[c] Subgroup of women not taking a supplement containing the corresponding nutrient.

[d] The unit for carotenoid/retinol biomarkers is µg/L, while the unit for self-reported carotenoid/retinol intake is µg/day.

[e] The unit for tocopherol biomarkers is mg/L, while the unit for self-reported tocopherol intake is mg/day.

[f] The unit for folate biomarkers is ng/mL, while the unit for self-reported folate intake is µg/day.

acid intakes (percentage of total fatty acids) estimated by each SFFQ, by 2 7DDRs, and by 4 ASA24s were approximately comparable. Compared with 7DDR measures for carotenoids, retinol, α-tocopherol, and folate, SFFQ2 tended to overestimate the absolute intakes. Mean intakes for these nutrients from the ASA24 were generally similar to estimates from 7DDRs.

Reproducibility

As Table 3 shows, data for all evaluated nutrients were moderately-to-highly reproducible by 2 SFFQs spaced 1 year apart (ICC = 0.40–0.75). The reproducibility of unadjusted nutrient intakes measured by 2 7DDRs with a 6-month interval was high for most nutrients (ICC = 0.24–0.78), with values generally being lower for carotenoids (ICC = 0.24–0.53). As expected, nutrient intakes estimated by ASA24 repeated every 3 months over a year had relatively low reproducibility (ICC = 0.11–0.37). For DLW, among 76 participants with 2 DLW measurements, the ICC was 0.72 if repeated after 6 months, 0.78 if repeated after 9 months, and 0.64 if repeated after 12 months; for all 76 women, the ICC for DLW-assessed total energy was 0.71 with an average interval of 8.8 months between replicated measurements. For absolute intakes of protein, sodium, and potassium, the reproducibility of biomarkers measured every 3 months over 1 year was low to moderate (ICC = 0.37–0.56). Most plasma nutrient concentrations were moderately-to-highly reproducible, except total saturated fatty acids (ICC = 0.21) and *trans*-fatty acids (ICC = 0.25). Almost all ICCs were reduced after adjustment for total energy intake energy and

several dietary and nondietary factors. The lower ICCs for most nutrients measured by ASA24 and biomarkers indicated large day-to-day within-person variation in nutrient intakes over the 1-year interval. Within-person and between-person coefficients of variation for repeated assessments of nutrients estimated by means of different methods are presented in Web Table 1.

Validity assessed by ranking of individuals according to intake

Utilizing recovery biomarkers as the comparison methods, the deattenuated Spearman correlation coefficients (r) for intakes of energy-adjusted protein, sodium, and potassium were similar among averaged ASA24s ($r = 0.54$, $r = 0.36$, and $r = 0.52$), the paper SFFQ2 ($r = 0.46$, $r = 0.48$, and $r = 0.49$), and the WebFFQ ($r = 0.46$, $r = 0.46$, and $r = 0.42$). The correlations of the averaged 7DDRs with biomarkers ($r = 0.67$, $r = 0.54$, and $r = 0.64$) were higher than those for the SFFQs and ASA24s (Figure 2 and Web Table 2). The correlations were lower without adjustment for energy and within-person variation (Web Table 2). For energy intake, we observed low correlations for both averaged ASA24 ($r = 0.20$) and SFFQ2 ($r = 0.11$) and modest correlation for 7DDR ($r = 0.46$). Correlations with biomarkers were slightly and nonsignificantly higher for the average of SFFQ1 and SFFQ2 than for SFFQ2 alone. Likewise, correlations were only slightly and nonsignificantly higher for the average of 2 7DDRs versus 1 7DDR (Web Table 3).

Correlations between nutrient intakes assessed via the 3 self-reporting methods and measured blood concentrations varied
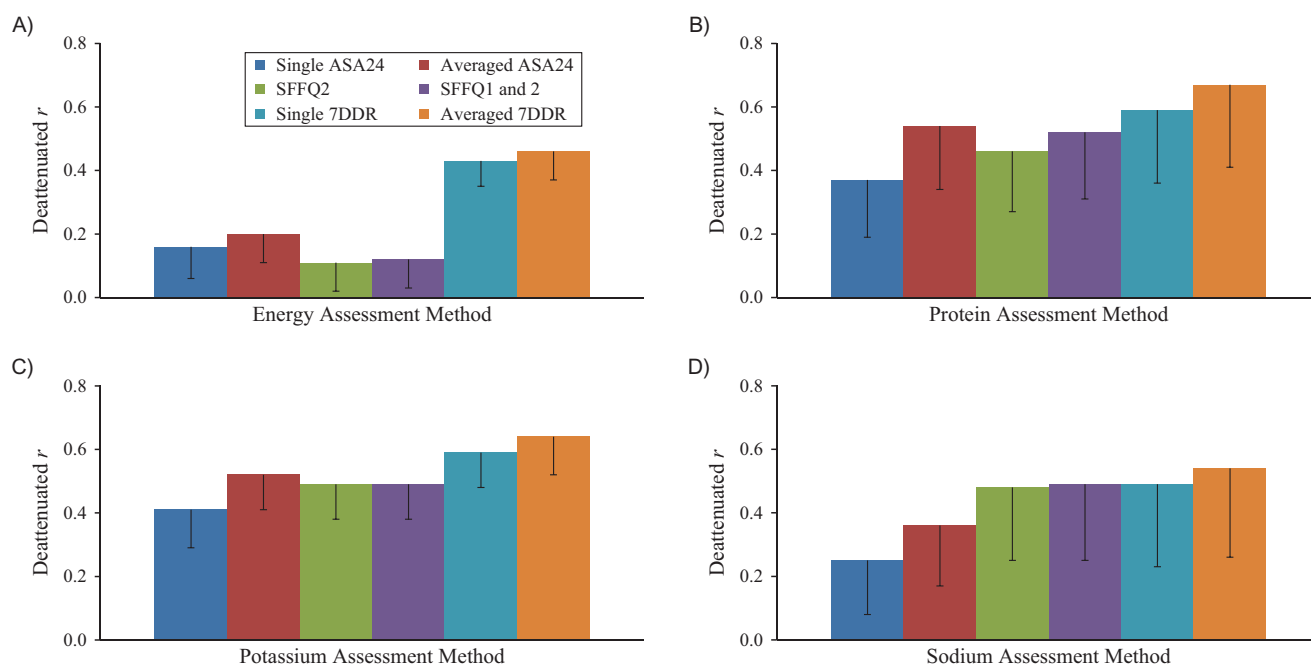
Figure 2. Deattenuated and energy-adjusted Spearman correlation coefficients for correlations between dietary intakes estimated from a semi-quantitative food frequency questionnaire (SFFQ), the Automated Self-Administered 24-Hour Dietary Recall (ASA24), 7-day dietary records (7DDRs), and urinary recovery biomarkers, Women's Lifestyle Validation Study, United States, 2010–2012. A) Energy (kcal/day); B) protein (percentage of energy); C) potassium (mg/1,000 kcal); D) sodium (mg/1,000 kcal). Correlation coefficients and corresponding *P* values for pairwise comparisons are presented in Web Tables 2 and 3. Bars, lower 95% confidence limit.
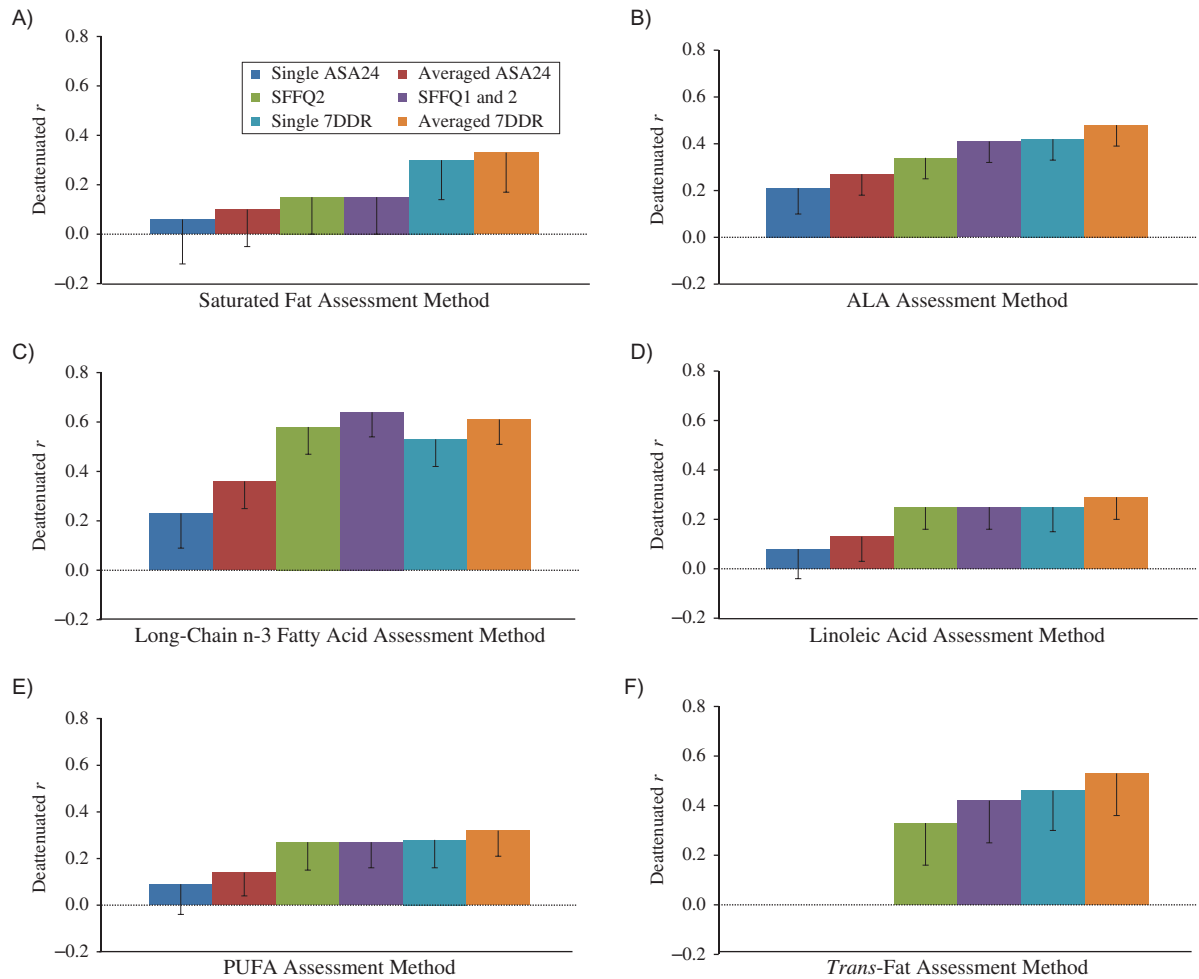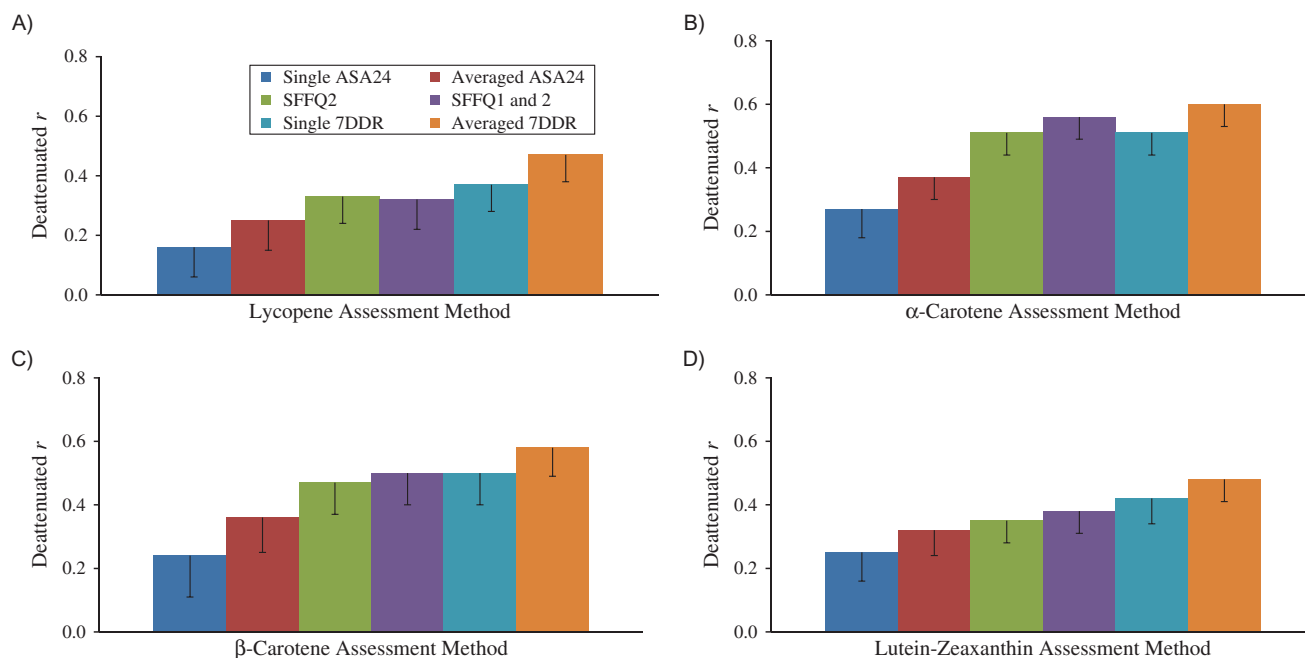
Figure 3. Deattenuated and adjusted Spearman correlation coefficients for correlations between fatty acid intakes (percentage of total fatty acids) estimated from a semiquantitative food frequency questionnaire (SFFQ), 7-day dietary records (7DDRs), the Automated Self-Administered 24-Hour Dietary Recall (ASA24), and plasma biomarkers, Women's Lifestyle Validation Study, United States, 2010–2012. A) Saturated fat; B) α-linolenic acid (ALA); C) long-chain n-3 fatty acids; D) linoleic acid; E) polyunsaturated fatty acids (PUFAs); F) *trans*-fat. Correlation coefficients and corresponding *P* values for pairwise comparisons are presented in Web Tables 4 and 6. Bars, lower 95% confidence limit.

from weak to strong after adjustment for energy and several dietary and nondietary factors and correction for random within-person variation in the biomarkers (Figures 3–5, Web Tables 4 and 5). Notably, after adjustment for energy intake, the additional adjustments for other dietary and nondietary factors made little difference in the correlation coefficients. Because plasma levels of saturated fatty acids, monounsaturated fatty acids, and retinol are largely not determined by intake, as expected, Spearman correlation coefficients for correlations between self-reported intakes of these nutrients and corresponding plasma levels were generally low for all self-reporting methods, even after correction for the random within-person variations in the comparison method. The deattenuated and adjusted correlations between averaged ASA24 records and biomarkers were 0.10 for saturated fat, 0.07 for monounsaturated fatty acids, 0.27 for α-linolenic acid, 0.36 for long-chain n-3 fatty acids, 0.13 for linoleic acid, 0.14 for polyunsaturated fatty acids, and 0.25–0.45 for most carotenoids. Among women who did not take supplements, averaged

ASA24s performed similarly to SFFQ2 in estimating folate intake (Figure 5). For most nutrients, compared with averaged ASA24s, the SFFQ2 had relatively higher correlations with biomarkers, especially for long-chain n-3 fatty acids (*r* = 0.58). The WebFFQ and the paper SFFQ2 had similar correlations with biomarkers of fatty acids, carotenoids, retinol, tocopherol, and folate (Web Tables 4 and 5). Two SFFQs had validity generally similar to that of 1 7DDR; the validity of the averaged 7DDRs was generally slightly but nonsignificantly higher than that of a single 7DDR (Web Tables 6 and 7). The results for all specific fatty acids are shown in Web Tables 8–10.

DISCUSSION

Utilizing both recovery and concentration biomarkers, we examined the pairwise correlations between 3 self-reporting dietary assessment methods and corresponding biomarkers
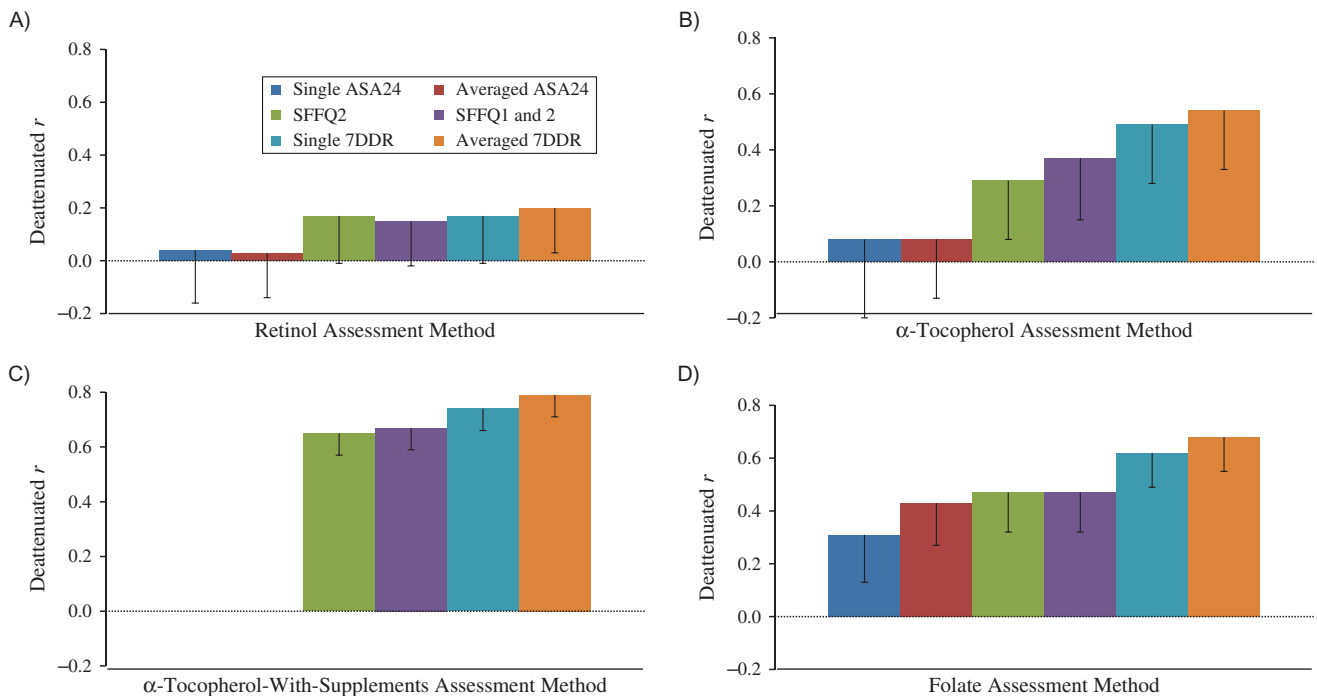
Figure 4. Deattenuated and adjusted Spearman correlation coefficients for correlations between carotenoid (μg/L or μg/day) intakes estimated from a semiquantitative food frequency questionnaire (SFFQ), 7-day dietary records (7DDRs), the Automated Self-Administered 24-Hour Dietary Recall (ASA24), and plasma biomarkers, Women's Lifestyle Validation Study, United States, 2010–2012. A) Lycopene; B) α-carotene; C) β-carotene; D) lutein-zeaxanthin. Correlation coefficients and corresponding *P* values for pairwise comparisons are presented in Web Tables 5 and 7. The unit for carotenoid biomarkers is μg/L, while the unit for self-reported carotenoid intake is μg/day. Bars, lower 95% confidence limit.

in the estimation of nutrient intakes, including intakes of total energy, protein, sodium, potassium, specific fatty acids, specific carotenoids, retinol, tocopherols, and folate. These data provide an indication of relative validity for assessing diet over a 15-month period; for recovery biomarkers, they also provide an indication of absolute validity, assuming the biomarkers have no systematic within-person error (person-specific bias). Judged by their correlations with the biomarkers, the average of approximately 3 ASA24s performed similarly to the SFFQ2 in estimating energy-adjusted protein, sodium, and potassium intakes. For absolute energy intake, we observed low correlations of both ASA24 and SFFQ with biomarkers. However, for almost all nutrients with concentration biomarkers, we observed a pattern that SFFQ2 had better performance than averaged ASA24s. In general, relative to biomarkers, single and averaged ASA24s had lower validity than SFFQ2; SFFQ2 had slightly lower validity than 1 7DDR, while the averaged SFFQs had validity similar to that of 1 7DDR; and the averaged 7DDRs had the highest validity. The average of 2 SFFQs had slightly higher validity than SFFQ2. Notably, 2 7DDRs performed only slightly better than 1 7DDR, suggesting that 2 7DDRs completed at an interval of about 6 months approaches the maximal potential validity for dietary assessment conducted by this method.

These study findings greatly extend those from comparisons of different dietary assessment methods using only recovery biomarkers (8–14), which exist for only a few nutrients. Recovery biomarkers of dietary intakes are considered an objective "gold standard" for validation of other dietary methods; however, they still have many limitations due to high day-to-day variations and very high costs. In a previous study, the Observing Protein and Energy Nutrition (OPEN) Study (14), dietary recalls had higher correlations with biomarkers than an FFQ, and it was suggested that validation studies using diet records or recalls as the comparison method may seriously overstate validity of an FFQ because of correlated errors. However, this conclusion was based on a single nutrient, protein, which behaves differently from most nutrients due to limited between-person variation. Furthermore, the relatively short time interval between repeated measurements of biomarkers would underestimate the validity of the SFFQs because the long-term variation in biomarker measurement was not well captured (16). In our study, for protein, sodium, and potassium density assessed by SFFQ2, the deattenuated correlations with their biomarkers were 0.46, 0.48, and 0.49 respectively, which were only slightly lower than the correlations between SFFQ intakes with 7DDRs reported in our previous study (*r* = 0.54, *r* = 0.53, and *r* = 0.65, respectively) (40). Our observed correlations were also similar to the reported results from a recent pooled analysis of 5 validation studies (*n* = 484, *n* = 263, *n* = 524, *n* = 544, and *n* = 450), which documented that the averaged correlations between FFQ and biomarkers were 0.43 for protein density, 0.32 for sodium density, and 0.47 for potassium density (8, 9). These findings suggest that SFFQ2 had moderate validity for many nutrients (except absolute energy intake) using recovery biomarkers as the comparison method and that its validity was not seriously overestimated when dietary records were used as the comparison method.

Figure 5. Deattenuated and adjusted Spearman correlation coefficients for correlations between retinol (µg/L or µg/day), α-tocopherol (mg/L or mg/day), and folate (ng/mL or µg/day) intakes estimated from a semiquantitative food frequency questionnaire (SFFQ), 7-day dietary records (7DDRs), the Automated Self-Administered 24-Hour Dietary Recall (ASA24), and plasma biomarkers, Women's Lifestyle Validation Study, United States, 2010–2012. A) Retinol; B) α-tocopherol; C) α-tocopherol with supplement use; D) folate. Correlation coefficients and corresponding *P* values for pairwise comparisons are presented in Web Tables 5 and 7. The unit for retinol biomarkers is µg/L, while the unit for self-reported retinol intake is µg/day. The unit for tocopherol biomarkers is mg/L, while the unit for self-reported tocopherol intake is mg/day. The unit for folate biomarkers is ng/mL, while the unit for self-reported folate intake is µg/day. Bars, lower 95% confidence limit.

Our study observed weak correlations for saturated fatty acid, monounsaturated fatty acids, and several other endogenously synthesized fatty acids, whereas fatty acids of mainly exogenous origin had higher correlations with biomarkers for SFFQ and 7DDRs. Previous studies also found modest-to-strong correlations between n-3 fatty acid, *trans*-fat, and linoleic acid intakes measured from an FFQ and biomarkers measured in plasma (52), erythrocytes (53), or adipose tissue (18, 27, 54). For carotenoids, α-tocopherol, and folate, we observed moderate-to-high correlations between the SFFQ and biomarkers (*r* ranged from 0.33 for lycopene to 0.65 for α-tocopherol), similar to most existing studies (6, 26). We also observed similar performance of the Web-based SFFQ and the paper version in comparison with both recovery and concentration biomarkers. Our study found that multiple ASA24s tended to have lower correlations with biomarkers for carotenoids and fatty acids, compared with estimates by SFFQs, presumably due to large day-to-day variation in intakes of these nutrients. For example, long-chain n-3 fatty acids are mainly derived from fish, which is an episodically consumed food, and a single ASA24 works poorly in assessing the intake of long-chain n-3 fatty acids. The performance of the average of 3 days of ASA24 was improved but was still not as good as a single SFFQ2. This limitation of dietary recalls likely also applies for many episodically consumed foods, such as specific vegetables, fruits, legumes, and nuts (55).

Among the intake methods used in this study, the 2 7DDRs had the strongest correlations with biomarkers of intake, presumably because of a total of 14 days of open-ended, weighed, and recorded intake and the detailed coding of foods. The high costs and participant burden of this method preclude its use in large studies, particularly if repeated assessments over time are needed. In addition, the 2 7DDRs should not be completed too close together in time, in order to capture seasonal variability in diet. Multiple ASA24s did not perform as well as the paper and Web SFFQs.

Concentration biomarkers are subject to many potential errors when used as the standard for dietary assessment, including individual differences in nutrient absorption and metabolism, endogenous synthesis, homeostatic control mechanisms, interactions with other nutrients, turnover in the targeted tissue, day-to-day or seasonal variations in dietary intakes, and technical error associated with laboratory measurement. However, they are of value in assessing the relative validity of different dietary intake methods. The interpretation of these results could be enhanced by calibration of these biomarkers to absolute intake in controlled feeding studies. Further validation studies could incorporate additional dietary biomarkers and evaluate the changes in biomarker levels in response to changes in diet. Further studies are needed to address the effects of potential systematic within-person errors in the biomarkers when used as standards to assess validity.

To the best of our knowledge, this is the largest validation study using multiple dietary assessment measures and biomarkers to have been carried out to date. By design, all short-term dietary assessment methods were repeated at reasonable time intervals (at least 3 months apart) to capture the variation in diet over a period of 15 months, and we did not administer those methods closely in time (they were administered 1–5 weeks apart from each other and in random order) in order to avoid artificially high correlations. In addition, note that the biomarker assessments were still more proximal to the period assessed by the ASA24 than the period assessed by the SFFQs. This could potentially cause overestimation of ASA24 correlations with long-term true intake and underestimation of SFFQ correlations (41). Investigators should consider these temporal relationships when designing future validation studies (6). Our study also had some limitations. Study participants were registered female nurses from the Nurses' Health Study and Nurses' Health Study II cohorts. Thus, our results may not be generalizable to other populations, such as those with large proportions of racial/ethnic minorities, persons with diagnosed chronic diseases, or men. Further, differences in food composition tables used for the various dietary assessment methods may also contribute to differences between the 3 methods, although this difference was minimized in our study because the nutrient databases used to analyze the 3 different dietary methods were partly based on US Department of Agriculture data.

In conclusion, our findings document that the SFFQ and dietary records, especially 2-week averages, provide reasonably valid measurements of a wide variety of energy-adjusted nutrient intakes. In general, multiple days of weighed diet records provide an optimal assessment of the dietary factors evaluated in this study, but for most nutrients, the validity of the SFFQ was only modestly less than that of the diet record. Because of its far lower cost and participant burden, the SFFQ will continue to play an important role in studies of diet and health. The optimal number of 24-hour recalls needed for measures to be comparable to those from an SFFQ requires further evaluation, but a small number of days will be inadequate for assessment of some nutrients.

ACKNOWLEDGMENTS

REFERENCES

1. Willett WC, ed. *Nutritional Epidemiology*. 3rd ed. New York, NY: Oxford University Press; 2012.
2. Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute. Dietary Assessment Calibration/Validation Register: studies and their associated publications. https://epi.grants.cancer.gov/dacv/. Updated November 22, 2016. Accessed August 30, 2017.
3. Van Dam RM, Hunter D. Biochemical indicators of dietary intake. In: Willett WC, ed. *Nutritional Epidemiology*. 3rd ed. New York, NY: Oxford University Press; 2012:150–212.
4. Bingham SA. Biomarkers in nutritional epidemiology. *Public Health Nutr*. 2002;5(6A):821–827.
5. Shai I, Rosner BA, Shahar DR, et al. Dietary evaluation and attenuation of relative risk: multiple comparisons between blood and urinary biomarkers, food frequency, and 24-hour recall questionnaires: the DEARR study. *J Nutr*. 2005;135(3):573–579.
6. Willett WC, Lenart EB. Reproducibility and validity of food-frequency questionnaires. In: Willett WC, ed. *Nutritional Epidemiology*. 3rd ed. New York, NY: Oxford University Press; 2012:96–141.
7. Ocke MC, Kaaks RJ. Biochemical markers as an additional measurement in dietary validity studies: application of the method of triads with examples from the European Prospective Investigation into Cancer and Nutrition. *Am J Clin Nutr*. 1997;65(4 suppl):1240S–1245S.
8. 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(2):172–188.
9. Freedman LS, Commins JM, Moler JE, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. *Am J Epidemiol*. 2015;181(7):473–487.
10. Prentice RL, Mossavar-Rahmani Y, Huang Y, et al. Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. *Am J Epidemiol*. 2011;174(5):591–603.
11. 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 whites: the Energetics Study. *J Am Diet Assoc*. 2010;110(6):857–864.

12. Neuhouser ML, Tinker L, Shaw PA, et al. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol*. 2008;167(10): 1247–1259.

13. Moshfegh AJ, Rhodes DG, Baer DJ, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *Am J Clin Nutr*. 2008;88(2): 324–332.

14. Subar AF, Kipnis V, Troiano RP, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN Study. *Am J Epidemiol*. 2003;158(1):1–13.

15. Schatzkin A, Kipnis V, Carroll RJ, 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(6):1054–1062.

16. Willett W. Invited commentary: OPEN questions. *Am J Epidemiol*. 2003;158(1):22–24.

17. Preis SR, Spiegelman D, Zhao BB, et al. Application of a repeat-measure biomarker measurement error model to 2 validation studies: examination of the effect of within-person variation in biomarker measurements. *Am J Epidemiol*. 2011; 173(6):683–694.

18. Hunter DJ, Rimm EB, Sacks FM, et al. Comparison of measures of fatty acid intake by subcutaneous fat aspirate, food frequency questionnaire, and diet records in a free-living population of US men. *Am J Epidemiol*. 1992;135(4):418–427.

19. Garland M, Sacks FM, Colditz GA, et al. The relationship between dietary intake and adipose tissue composition of selected fatty acids in US women. *Am J Clin Nutr*. 1998;67(1): 25–30.

20. Ascherio A, Stampfer MJ, Colditz GA, et al. Correlations of vitamin A and E intakes with the plasma concentrations of carotenoids and tocopherols among American men and women. *J Nutr*. 1992;122(9):1792–1801.

21. Michaud DS, Giovannucci EL, Ascherio A, et al. Associations of plasma carotenoid concentrations and dietary intake of specific carotenoids in samples of two prospective cohort studies using a new carotenoid database. *Cancer Epidemiol Biomarkers Prev*. 1998;7(4):283–290.

22. Giovannucci E, Stampfer MJ, Colditz GA, et al. Folate, methionine, and alcohol intake and risk of colorectal adenoma. *J Natl Cancer Inst*. 1993;85(11):875–884.

23. Zhang SM, Willett WC, Selhub J, et al. Plasma folate, vitamin B6, vitamin B12, and homocysteine and risk of breast cancer. *J Natl Cancer Inst*. 2003;95:373–380.

24. Feskanich D, Willett WC, Colditz GA. Calcium, vitamin D, milk consumption, and hip fractures: a prospective study among postmenopausal women. *Am J Clin Nutr*. 2003;77(2): 504–511.

25. Sun Q, Ma J, Campos H, et al. Comparison between plasma and erythrocyte fatty acid content as biomarkers of fatty acid intake in US women. *Am J Clin Nutr*. 2007;86(1):74–81.

26. Park JY, Vollset SE, Melse-Boonstra A, et al. Dietary intake and biological measurement of folate: a qualitative review of validation studies. *Mol Nutr Food Res*. 2013;57(4): 562–581.

27. Baylin A, Campos H. The use of fatty acid biomarkers to reflect dietary intake. *Curr Opin Lipidol*. 2006;17(1):22–27.

28. Hedrick VE, Dietrich AM, Estabrooks PA, et al. Dietary biomarkers: advances, limitations and future directions. *Nutr J*. 2012;11:109.

29. Walker AM, Blettner M. Comparing imperfect measures of exposure. *Am J Epidemiol*. 1985;121(6):783–790.

30. Kirkpatrick SI, Subar AF, Douglass D, et al. Performance of the Automated Self-Administered 24-Hour Recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. *Am J Clin Nutr*. 2014;100(1):233–240.

31. 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(4):340–347.

32. Rimm EB, Giovannucci EL, Stampfer MJ, et al. Reproducibility and validity of a expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol*. 1992;135(10):1114–1126.

33. Krisberg K. Findings from Nurses' Health Study benefit women's health: researchers recruiting for third round. *The Nation's Health*. 2013;43(9):1–12.

34. Lawson CC, Johnson CY, Chavarro JE, et al. Work schedule and physically demanding work in relation to menstrual function: the Nurses' Health Study 3. *Scand J Work Environ Health*. 2015;41(2):194–203.

35. Casey PH, Goolsby SL, Lensing SY, et al. The use of telephone interview methodology to obtain 24-hour dietary recalls. *J Am Diet Assoc*. 1999;99(11):1406–1411.

36. National Cancer Institute. *White Paper: Biomarker-Based Validation of Diet and Physical Activity Assessment Tools in Existing Cohorts*. Bethesda, MD: National Cancer Institute; 2013. https://healthcaredelivery.cancer.gov/tools/biomarker_validation_in_cohorts.pdf. Accessed February 13, 2018.

37. Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J Womens Health*. 1997;6(1):49–62.

38. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer*. 2005;5(5):388–396.

39. Cho E, Chen WY, Hunter DJ, et al. Red meat intake and risk of breast cancer among premenopausal women. *Arch Intern Med*. 2006;166(20):2253–2259.

40. Yuan C, Spiegelman D, Rimm EB, et al. Validity of a dietary questionnaire assessed by comparison with multiple weighed dietary records or 24-hour recalls. *Am J Epidemiol*. 2017; 185(7):570–584.

41. Freedman LS, Midthune D, Carroll RJ, et al. Application of a new statistical model for measurement error to the evaluation of dietary self-report instruments. *Epidemiology*. 2015;26(6): 925–933.

42. Schakel SF, Sievert YA, Buzzard IM. Sources of data for developing and maintaining a nutrient database. *J Am Diet Assoc*. 1988;88(10):1268–1271.

43. Schakel SF, Buzzard IM, Gebhardt SE. Procedures for estimating nutrient values for food composition databases. *J Food Compost Anal*. 1997;10(2):102–114.

44. Bingham SA, Cummings JH. Urine nitrogen as an independent validatory measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet. *Am J Clin Nutr*. 1985;42(6):1276–1289.

45. Holbrook JT, Patterson KY, Bodner JE, et al. Sodium and potassium intake and balance in adults consuming self-selected diets. *Am J Clin Nutr*. 1984;40(4):786–793.

46. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr*. 1997;65(4 suppl):1220S–1228S.

47. Beaton GH, Milner J, McGuire V, et al. Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. *Am J Clin Nutr*. 1983;37(6):986–995.

48. Rosner B, Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. *Am J Epidemiol*. 1988; 127(2):377–386.
49. Perisic I, Rosner B. Comparisons of measures of interclass correlations: the general case of unequal group size. *Stat Med*. 1999;18(12):1451–1466.
50. Rosner B, Glynn RJ. Interval estimation for rank correlation coefficients based on the probit transformation with extension to measurement error correction of correlated ranked data. *Stat Med*. 2007;26(3):633–646.
51. Chavarro JE, Rosner BA, Sampson L, et al. Validity of adolescent diet recall 48 years later. *Am J Epidemiol*. 2009; 170(12):1563–1570.
52. Ma J, Folsom AR, Shahar E, et al. Plasma fatty acid composition as an indicator of habitual dietary fat intake in middle-aged adults. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. *Am J Clin Nutr*. 1995;62(3):564–571.
53. Sun Q, Ma J, Campos H, et al. Comparison between plasma and erythrocyte fatty acid content as biomarkers of fatty acid intake in US women. *Am J Clin Nutr*. 2007;86(1):74–81.
54. Baylin A, Kabagambe EK, Siles X, et al. Adipose tissue biomarkers of fatty acid intake. *Am J Clin Nutr*. 2002;76(4): 750–757.
55. 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(4):340–347.