Published in final edited form as:
Curr Atheroscler Rep. 2013 September ; 15(9): . doi:10.1007/s11883-013-0353-5.

Calibration Of Self-Reported Dietary Measures Using Biomarkers: An Approach To Enhancing Nutritional Epidemiology Reliability

Ross L. Prentice, PhD[1], **Lesley F. Tinker, PhD**[2], **Ying Huang, PhD**[3], and **Marian L. Neuhouser, PhD**[4]
Fred Hutchinson Cancer Research Center, Division of Public Health Sciences, 1100 Fairview Avenue North, M3-A410, POB 19024, Seattle, WA 98109-1024

Abstract

Reports from nutritional epidemiology studies lack reliability if based solely on self-reported dietary consumption estimates. Consumption biomarkers are available for some components of diet. These can be collected in subsets of study cohorts, along with corresponding self-report assessments. Linear regression of (log-transformed) biomarker values on corresponding self-report values and other pertinent study subject characteristics yields calibration equations for dietary consumption, from which calibrated consumption estimates can be calculated throughout study cohorts. Nutritional epidemiology disease association studies of enhanced reliability can be expected from analyses that relate disease risk to calibrated consumption estimates. Applications to the study of energy and protein consumption in relation to cardiovascular diseases, type 2 diabetes, and cancer in the Women's Health Initiative will be briefly summarized. Also, challenges related to variables that may either mediate or confound associations of interest will be described, along with the need for longitudinal biomarker and self-report data, and the need for additional nutritional biomarkers development.

Keywords

biomarker; calibration; cardiovascular disease; epidemiology; nutrition; measurement error

Introduction

In spite of substantial public interest, and great public health importance, there are few convincing or consistent disease association results from several decades of intensive nutritional epidemiology research. Positive associations between obesity and the risk of prominent cardiovascular diseases and cancers, among other chronic diseases, are well-established, but reviews of the world literature find few dietary associations that are regarded as convincing or probable [1–3].

Corresponding Author: Ross L. Prentice.
[1]Phone: 206-667-4264; Fax: 206-667-4142; rprentic@whi.org
[2]Phone: 206-667-6894; Fax: 206-667-4142; ltinker@whi.org
[3]Phone: 206-667-4198; Fax: 206-667-4378; yhuang@fhcrc.org
[4]Phone: 206-667-4797; Fax: 206-667-5977; mneuhous@fhcrc.org

Conflict of Interest:
Ross L. Prentice, Lesley F. Tinker, Ying Huang, and Marian L. Neuhouser declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent
This article does not contain any studies with human or animal subjects performed by any of the authors.

Most nutritional epidemiology studies are observational, with cohort studies providing the context for most reports in recent years. As such, these studies need to contend with the usual observational study challenges of confounding and unbiased outcome ascertainment. Additionally, they need to address the thorny and potentially dominating issue of dietary assessment measurement error.

Dietary assessment in epidemiologic cohort study contexts has relied mainly on a food frequency questionnaire (FFQ) assessment procedure, which is quite practical for use in large cohorts due to its usual self-administered and economical machine-readable features. Nutritional epidemiology reports that attempt to address the measurement issue have sometimes included a second dietary assessment procedure, such as one or more 24-hour dietary recalls (24HRs) or multi-day food records, from which (log-transformed) nutrient or food consumption estimates are assumed to adhere to a classical measurement model. The classical model assumes additive error that is unrelated to the targeted consumption, unrelated to other study subject characteristics, and independent of the corresponding FFQ measurement error. Although this approach may frequently lead to some reduction in bias due to measurement error [4], the assumptions just listed for a second dietary assessment approach do not appear to hold, even approximately, for available self-report procedures for energy and protein assessment [5].

On the other hand, for some dietary factors, there are established consumption biomarkers for which these measurement modeling assumptions are quite plausible, at least for short-term consumption. This paper summarizes some experiences to date in the use of total energy and protein consumption biomarkers in Women's Health Initiative (WHI) cohorts, and sets out some study design, data collection, and biomarker development needs to support the use of a biomarker calibration approach more generally in nutritional epidemiology research.

Before doing so, it may be useful to comment briefly on the role and potential of randomized trial designs in nutritional epidemiology. Randomized, controlled intervention trials have the considerable advantage of avoiding confounding by all pre-randomization factors, whether recognized as such or not. Such trials typically also provide a context for unbiased outcome ascertainment. Importantly, the validity of intention-to-treat comparisons in randomized trials does not depend on dietary assessment, though some form of dietary assessment for the randomized groups as a whole is needed to document that a comparison of acceptable statistical power has taken place. However, full-scale randomized controlled intervention trials, such as the Dietary Modification (DM) Trial among 48,835 postmenopausal women in the WHI [6] are logistically difficult and expensive, and may require a lengthy follow-up period for the hypothesized intervention effects on disease risk to have time to emerge. Hence, only a few full-scale intervention trials can be afforded, and any such trial requires a careful developmental process. On the other hand, clear results from randomized intervention trials, such as the major reduction in diabetes incidence among persons assigned to the diet and physical activity intervention in the Diabetes Prevention Program [7], can have a profound effect on nutritional epidemiology research and knowledge. Also, results from intermediate outcome randomized controlled trials of practical size, duration and cost may be able to be combined with observational data relating intermediate outcomes to disease risk to provide a novel strategy within an overall nutritional epidemiology research agenda. For example, heart disease related intermediate outcomes trials could include blood pressure, serum lipids, coagulation factors, inflammatory factors , and even high dimension data such as metabolomic profiles.

Biomarker Calibration Approach to Nutritional Epidemiology

The biomarker calibration approach to nutritional epidemiology research relies on the availability of objective markers that plausibly adhere to a classical measurement model for the dietary variables under study. Objective measures of this type mostly arise from urinary recovery as nutrients are metabolized and excreted [8], but could also arise from measures in blood or other body fluids or tissues. Prominent examples include a doubly-labeled water (DLW) biomarker of total energy expenditure [9] and a urinary nitrogen (UN) biomarker of protein intake [10]. These recovered nutrients reflect short-term consumption among weight stable persons. Other biomarkers include urinary potassium and sodium which, like protein, can be estimated from 24-hour urine collections. The energy biomarker is more specialized and expensive, involving the consumption of a few ounces of DLW that includes enriched doses of the naturally occurring (safe to consume) stable isotopes O^{18} and deuterium, with recovery of metabolites reflecting oxygen and carbon dioxide expenditure over the ensuing two weeks [9].

Two nutritional biomarker sub-studies have been conducted within WHI cohorts, and a third that aims to develop novel biomarkers is nearing completion. The first, referred to as the Nutrient Biomarker Study (NBS), included DLW and UN biomarkers and concurrent FFQs among 544 women in the WHI DM trial, a trial that examined the effects of a low fat, diet with increased vegetables, fruits and grains on cancer and cardiovascular disease (CVD) outcomes, with 50% of women from the low-fat dietary pattern intervention group, and 50% from the usual diet comparison group [11]. The NBS took place during 2004–2006 at 12 of the 40 clinical centers that participated in the WHI. A 20% reliability subsample completed the entire protocol about six months after the initial application.

The second biomarker study, conducted during 2007–2010 and termed the Nutrition and Physical Activity Assessment Study (NPAAS), involved these same biomarkers, along with concurrent FFQs, as well as 4-day food records (4DFRs) and three 24-hour dietary recalls, among 450 women at nine WHI clinical centers who participated in the WHI Observational Study (OS). The OS is a prospective cohort study among 93,676 postmenopausal women drawn from essentially the same catchment area as the WHI clinical trials, with much commonality in epidemiologic and clinical data collection [6]. The NPAAS protocol also included indirect calorimetry for resting energy expenditure assessment, and physical activity self-report data collection using a 7-day physical activity recall, the Arizona Activity Frequency Questionnaire, and a WHI questionnaire-based physical activity assessment [5, 12]. As with the NBS, the entire protocol was repeated about six months after the initial application by a 20% subsample of the participating women.

The biomarker calibration approach assumes a classical measurement model for the objective measure, while allowing a more flexible model for the corresponding self-report. For example, the biomarker assessment W may be the logarithm of DLW average daily total energy consumption over the two-week period of a biomarker study, while the targeted dietary consumption Z is the logarithm of total average daily energy consumption over a longer time period (e.g., one-year period from three months before to nine months after the biomarker study), and Q is the logarithm of self-reported average daily energy consumption over a closely related time period.

Biomarker values are assumed to arise from a classical measurement model

$$W = Z + u$$

with error u that is independent of Z, and independent of all relevant study subject characteristics V = (v_1, v_2, …). For example, V may include factors needed to avoid confounding in studying the association between Z and a study disease, or factors related to the magnitude of the self-report measurement error. The corresponding self-report Q is allowed to have a biased 'target', Z^*, where

$$Z^* = a_0 + a_1 Z + a_2{}^{\mathrm{T}} V$$

so that

$$Q = Z^* + e,$$

where the error term e is assumed to be independent of Z and u, given V.

A joint normality assumption for (Z, e, V) then gives expectation of the form

$$E(Z|Q, V) = b_0 + b_1 Q + b_2{}^{\mathrm{T}} V,$$

from which the biomarker model assumptions give

$$E(W|Q, V) = b_0 + b_1 Q + b_2{}^{\mathrm{T}} V$$

It follows that one can obtain estimates of the underlying dietary consumption Z given (Q, V) under these modeling assumptions, by linear regression of the biomarker (W) on the corresponding self-report (Q) and other pertinent study subject characteristics. The resulting linear equation

$$\hat{Z} = \hat{b}_0 + \hat{b}_1 Q + \hat{b}_2{}^{T} V$$

can be used to generate 'calibrated' consumption estimates $\hat{Z}$ from corresponding (Q, V) values throughout the larger study cohort. This type of calibration equation can be expected to correct systematic biases related to V in the self-report Q. Also, a good calibration equation should be able to 'recover' a substantial fraction of the variation in Z in the study population. For this latter purpose, V may include study subject characteristics that enhance the percent of variation explained by the equation, even if not needed for avoidance of bias due to the self-report measurement error properties or for avoidance of confounding bias. Note that the above development can readily be relaxed to reduce the normality assumption to that of a joint normal distribution for (Z, e) given V.

The biomarker W typically also incorporates measurement error (u), both technical error in relation to consumption in the short time period targeted by the biomarker, and temporal error relative to the larger time period used to define Z. Reliability subsample data can provide a means for adjusting for temporal variation, at least over a moderate time period. Note, however, if Z is defined in terms of consumption over the years or decades that may be pertinent to disease risk, then longitudinal biomarkers and self-report data over a longer time period will typically be needed for reliable disease association estimation.

The concept behind the use of calibrated consumption estimates derives from the so-called regression calibration approach to hazard ratio parameter estimation [13–15] in disease

association analyses. Specifically, a hazard ratio (Cox) model for a time-to-disease occurrence outcome (T) may be written as

$$\lambda(t;Z,V)=\lambda_0(t)\exp(Z\alpha_1+V^T\alpha_2)$$

with α_1 the hazard ratio parameter of interest, and V comprised of factors that may, if omitted, confound the association between Z and disease risk. Since the self-report Q, rather than Z, is observed, a hazard rate model is induced that, for a relatively rare disease is, to a good approximation [13],

$$\lambda^*(t;Q,V)=\lambda^*{}_0(t)\exp(E(Z|Q,V)\alpha_1+V^T\alpha_2).$$

The calibrated values $\hat{Z}$ then provide an estimator of the conditional expectation E (Z | Q, V), leading to standard Cox regression estimates for α_1 and α_2 with modeled covariate $\hat{Z}$ and V. Because of uncertainty in the calibration equation coefficients, a non-standard variance estimator of 'sandwich' [14] or bootstrap form, is advisable, though this will typically differ little from the usual (partial likelihood) variance estimator from Cox regression if the biomarker subsample is fairly large.

The above concept implies that factors that may confound the association between Z and disease risk should be considered for inclusion in calibration equation development. Decisions concerning which variables, V, are needed in the hazard ratio model λ (t; Z, V) can be challenging, particularly if it is unclear whether a study subject characteristic should be regarded as a confounder or a mediator of the diet and disease association, or both. This challenge would exist even if Z could be accurately measured in all members of the study cohort, and full resolution in an observation setting would require the collection of suitably rich longitudinal data on Z and V to examine temporal relationships. These challenges are exacerbated in the dietary measurement error context, particularly if the components of V under consideration as potential confounders or mediators are also important elements of consumption calibration. The example of average daily energy consumption in relation to CVD, diabetes, or cancer risk, with body mass index (BMI) as potential confounder, mediator, and calibrator will be considered in the next section to illustrate these issues.

Calibration Equations and Disease Associations in the Women's Health Initiative

Most applications of the biomarker calibration approach in WHI have involved energy and protein consumption, and their protein density ratio. These include association analyses with site-specific cancers [16], cardiovascular diseases [17], diabetes [18], frailty [19], and kidney function [20]. Other potential applications involve calibration of sodium and potassium consumption, and their ratio in relation to cardiovascular disease risk, and respiratory quotient calibration in an attempt to distinguish fat from carbohydrate utilization in relation to breast cancer or CVD risk.

To illustrate these approaches and analyses, Table 1, adapted from [5], shows a version of calibration equations from NPAAS for total daily energy (kilocalories), protein (grams) and protein density(percent energy from protein). Separate equations are shown when using FFQ, 4DFR, and 24HRs as sources of the self-report data. Beyond the self-report (Q), these equations include BMI and age at the time of the biomarker study, along with indicator variables for race/ethnicity. Note that the (log-transformed) self-report provides an explanation for only 3.8% (FFQ), 7.8% (4DFR), and 2.8% (24HR) of the log DLW

biomarker variation for total energy, but that age and race/ethnicity, and especially BMI, provide an explanation for considerably more of the biomarker variation, giving percent of variation (R^2) values in the 40–45% range with each of the self-reports. Our interest resides in calibrated estimates that capture the variation in Z, which could, as noted above, be defined as average daily energy consumption from three months before until nine months after a woman's beginning of NPAAS participation. Compared to this Z definition, the DLW measure (W) will include temporal measurement error as well as some technical measurement error. The 20% reliability subsample can be used to develop adjusted R^2 values that aim to estimate the percent of variation in Z, rather than in W, explained by the calibration equation. In fact, under the assumption that measurement errors at the two biomarker assessment times are independent, as may be a reasonable assumption with this 1-year Z definition, the adjusted R^2 values are obtained simply by dividing the unadjusted R^2 values by the sample correlation between the paired W values in the reliability subsample. Adjusted R^2 values in the 70–77% range suggest an excellent ability of the calibrated estimates to recapitulate the population variations in total energy consumption.

Calibrated protein consumption exhibited rather similar patterns with a somewhat weaker dependence on BMI. The adjusted R^2 values are again quite impressive, especially when 4DFRs or 24HRs are used for dietary assessment. Adjusted R^2 values are similarly large for protein density calibration using any of the three dietary assessment procedures. It is not clear that BMI is needed for protein density calibration.

Measurement error correlations were also estimated among the three dietary assessment procedures. The estimated FFQ/4DFR, FFQ/24HR, and 4DFR/24HR measurement error correlations based on logtransformed data were respectively 0.30, 0.30, and 0.50 for energy; 0.35, 0.35, and 0.27 for protein; and 0.38, 0.38, and 0.40 for protein density, with each much greater than zero [5]. These correlations argue against the use of one self-report to calibrate another, at least for a fully satisfactory approach to dietary measurement error correction.

More detailed calibration equations than are displayed in Table 1 may be needed in disease association analyses, depending on whether modeled confounding factors also contribute to calibration equations. For example, analyses of energy and protein in relation to cardiovascular disease in WHI cohorts [17] included age, ethnicity, education, household income, cigarette smoking status, and an estimate of recreational physical activity for confounding control. Calibration equations from NBS [11, Table 4] considered each of these listed variables, along with BMI and an indicator for intervention group randomization assignment in the DM trial. Table 2, adapted from [17] and [21], shows estimated coronary heart disease and stroke hazard ratios (HRs) and corresponding 95% confidence intervals associated with a 20% increment in energy, protein, or protein density from a Cox model analysis under which the logarithm of dietary consumption relates linearly to the log-hazard ratio. Uncalibrated FFQ energy is not associated with CHD or stroke, whereas calibrated energy is positively associated with CHD, and borderline inversely associated with stroke in analyses that do not include BMI in the disease risk model. When BMI is added to the disease risk model, the positive energy association with CHD disappears, whereas the inverse association with stroke becomes more apparent. If BMI is not a confounder of the energy and CHD risk association, as may be plausible if Z is defined as (log-transformed) long-term average daily energy intake over preceding years or decades, then Table 2 implies a positive association between total energy consumption and disease risk that appears to be fully mediated by body fat deposition over time. On the other hand, if BMI is a pure confounder, as may be plausible if Z is defined to reflect short-term energy consumption, then the Table 2 analyses do not suggest any energy association with CHD risk.

The right three columns in Table 2, from [21], provide further insight into these interpretations by beginning with a Cox model that may include short-term or longer-term average daily energy consumption and that also includes BMI, and then induces a hazard rate model for calibrated energy by taking expectation over the BMI distribution, given the other modeled variables. The HR estimate associated with average daily energy consumption then turns out to depend rather crucially on the log (DLW) energy measurement error correlations between the initial and reliability components in NBS. If these correlations are large and positive (e.g., 0.5), as may be the case if Z represents long-term energy consumption, then HR estimates like the previously described calibrated energy HRs without adjusting for BMI are obtained, whereas if such correlations are small (e.g., 0), as may be the case if Z represents short-term average energy consumption, then HRs like these adjusted for BMI are obtained. It seems clear that biomarker and self-report data over an extended time period are needed to accurately assess the energy consumption and CHD risk association. Regardless of the magnitude of any such positive association, however, these data suggest that any such association is rather fully mediated by body fat deposition over time.

The inverse association between energy consumption and stroke in Table 2 is intriguing, and is particularly evident when adjusting for BMI. The basis for this inverse association is unclear but it could, for example, reflect low energy consumption as a result of comorbidity, or high energy consumption due to relatively high physical activity. Similarly, total protein consumption may be inversely associated with both CHD and stroke, among women of a specified BMI, as may also be the case for protein density in relation to CHD. This later inverse association is also suggested by uncalibrated FFQ protein density, but with HR that is substantially attenuated toward the null. In summary, the use of biomarker calibrated energy and protein consumption estimates reveals cardiovascular disease associations that are mostly not evident without calibration, and suggests designs for future studies to facilitate interpretation of such associations.

Also, calibrated energy relates positively and strongly to type 2 diabetes risk in WHI cohorts, with an HR (95% CI of 2.41(2.06,2.82) for a 20% energy increment, that is reduced to 1.30 (0.96, 1.76) after controlling for BMI in the disease risk model [18].

Similarly, calibrated energy is positively associated with the risks of several site-specific (invasive) cancers in WHI cohorts [16]. Specifically, the HRs (95% CIs) for a 20% increment in energy were 1.24 (1.11, 1.38) for breast cancer, 1.35 (1.06, 1.71) for colon cancer, 1.83 (1.49, 2.25) for endometrium cancer, 1.47 (1.00, 2.16) for kidney cancer, and 1.18 (1.10, 1.27) for total invasive cancer , in analyses that do not adjust for BMI and therefore can be thought of in terms of association with long-term energy consumption[16]. These associations too were not apparent at all without biomarker calibration, and they largely disappeared upon controlling for BMI in the disease rate model.

Biomarker Development Needs

For the biomarker calibration approach to impact observational nutritional epidemiology in a comprehensive fashion, it will be necessary to develop biomarkers that adhere to a classical measurement model, with error terms that are independent of corresponding self-report error terms, for many nutrients and dietary components. Biomarkers such as DLW for energy and UN for protein have mostly been developed in small-scale studies with participants living for some days in specialized metabolic facilities where pertinent consumptions were closely controlled. A rather different study design is being employed in the third WHI nutritional biomarker study. This ‘NPAAS feeding study’, being conducted among 150 WHI enrollees in Seattle WA, is nearing completion. Participating women are provided food and drink that

approximates their usual dietary patterns so that urine and blood measures can stabilize quickly, by the end of a two-week feeding period, and so that most consumption variation in the study population will be preserved. Only foods having well-characterized nutrient composition are utilized. Linear regression of log-transformed provided nutrients/foods on corresponding urine and/or blood measures, along with other measured study subject characteristics, is the principal approach to biomarker development. Regressions that can 'recover' a substantial fraction (e.g., 50% or more) of the provided nutrient/food variation will provide novel biomarker candidates. If warranted, such biomarkers will then be determined using stored specimens from NBS and NPAAS for the calculation of corresponding calibrated consumption estimates throughout WHI cohorts, for use in novel disease association analyses. Feeding study investigators have several designated potential biomarkers (e.g., for sugars, whole grains, and meats), and will also utilize some higher-dimensional approaches using urine and blood metabolomic profiles.

Conclusions

Nutritional epidemiology is a most challenging research area, but also one of great public health importance, especially as the obesity epidemic and ensuing chronic disease risk continues to accelerate. A vigorous research program is needed that combines intervention trials of novel design and practical size with observational studies of the greatest possible reliability. Measurement error of a complex nature distinguishes nutritional epidemiology from many other areas of observational epidemiologic enquiry. For example, overweight and obese persons may underestimate their energy consumption by 30–50%, playing havoc with disease rate association estimation, unless suitably corrected. The ubiquitous FFQ gives energy consumption estimates that are essentially uncorrelated with BMI in WHI populations, even though DLW energy consumption correlates strongly (correlation of 0.8) with BMI.

A biomarker calibration approach offers potential for a fresh look at nutritional epidemiology associations to the extent that biomarkers are available for dietary exposures of interest. Biomarker studies of practical size and acceptable cost can be conducted within study cohorts to support biomarker-calibrated dietary association studies. Such studies will need to incorporate longitudinal biomarker and concurrent dietary self-report data collection to sort out such challenging topics as the relative confounding and mediating roles for BMI, in the context of total energy consumption and disease risk association analyses.

Many of these same comments apply to the assessment of activity-related energy expenditure [12], and its components, with the likelihood that accelerometers or other devices may prove valuable for objective assessments, either through cohort sub-studies in conjunction with corresponding self-report assessments, or through application to entire study cohorts.

In summary, the development of longitudinal biomarker and self-report data in cohort sub-studies, and the development of a comprehensive set of nutritional and physical activity biomarkers deserve a prominent place in the chronic disease population science, and the disease prevention, research agenda in upcoming years.

References

Papers of particular interest have been highlighted as:

• Of importance

•• Of major importance

1. WHO Technical Report 916. Geneva: World Health Organization; 2003. World Health Organization: Diet, nutrition and the prevention of chronic diseases: report of a joint WHO/FAO expert consultation; p. 88
2. Food, nutrition and the prevention of cancer: a global perspective. Washington, DC: American Institute for Cancer Research; 1997. World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) (1997).
3. World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Washington, DC: American Institute for Cancer Research; 2007.
4. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. J Natl Cancer Inst. 2011; 103:1–7.
5. Prentice RL, Mossavar-Rahmani Y, Huang Y, et al. Evaluation and comparison of food records, recalls and frequencies for energy and protein assessment using recovery biomarkers. Am J Epidemiol. 2011; 174:591–603. [PubMed: 21765003] **This paper provides a detailed account of calibration equation development for energy, protein, and protein/energy in the Women's Health Initiative Observational (prospective cohort) Study. Equations are developed using self-report data from food frequency questionnaires, four-day food records, and three 24-hour dietary recalls. Through comparison with doubly-labeled water and urinary nitrogen biomarkers of energy and protein consumption, the three self-report assessments are found to have measurement errors that are strongly positively related.
6. Women's Health Initiative Study Group. Design of the Women's Health Initiative Clinical Trial and Observational Study. Control Clin Trials. 1998; 19:61–109. [PubMed: 9492970]
7. Knowler W, Barrett-Connor E, Fowler S, et al. Reduction in the incidence of Type 2 Diabetes with lifestyle intervention or metformin. N Engl J Med. 2002; 346:393–403. [PubMed: 11832527]
8. Kaaks R, Ferrari P, Ciampi A, et al. Uses and limitations of random error correlations, in the validation of dietary questionnaire assessments. Pub Health Nutr. 2002; 5:969–976. [PubMed: 12638598]
9. Schoeller DA. Recent advances from application of doubly-labeled water to measurement of human energy expenditure. J Nutr. 1999; 129:1765–1768. [PubMed: 10498745]
10. Bingham SA. Urine nitrogen as a biomarker for the validation of dietary protein intake. J Nutr. 2003; 133:921S–924S. [PubMed: 12612177]
11. Neuhouser ML, Tinker L, Shaw PA, et al. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. Am J Epidemiol. 2008; 167:1247–1259. [PubMed: 18344516]
12. Neuhouser ML, Chongzhi D, Tinker LF, et al. Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. Am J Epidemiol. 2013; 177:576–585. [PubMed: 23436896] *This paper develops calibration equations for activity-related energy (AREE) expenditure based on three self-report assessment procedures. Self-report alone explains only a small fraction of biomarker AREE variation in Women's Health Initiative cohorts, but calibrated AREE explain substantially more.
13. Prentice RL. Covariate measurement errors and parameter estimation in a failure time regression model. Biometrika. 1982; 69:331–342.
14. Wang CY, Hsu L, Feng ZD, Prentice RL. Regression calibration in failure time regression with surrogate variables. Biometrics. 1997; 53:131–145. [PubMed: 9147589]
15. Carroll, RJ.; Ruppert, D.; Stefanski, LA.; Crainiceanu, CM. Measurement Error in Nonlinear Models, A Modern Perspective. 2nd Ed. Boca Raton, FL: Chapman and Hall/CRC; 2006.
16. Prentice RL, Shaw PA, Bingham SA, et al. Biomarker-calibrated energy and protein consumption and increased cancer risk among postmenopausal women. Am J Epidemiol. 2009; 169:977–989. [PubMed: 19258487]
17. Prentice RL, Huang Y, Kuller LH, et al. Biomarker-calibrated energy and protein consumption and cardiovascular disease risk among postmenopausal women. Epidemiology. 2011; 22:170–179. [PubMed: 21206366] *This paper reports calibrated energy consumption to be positively associated with coronary heart disease (CHD) incidence, and inversely associated with ischemic

stroke incidence in Women's Health Initiative cohorts. The positive CHD association appeared to be mediated by body fat deposition over time, as assessed by body mass index.

18. Tinker LF, Sarto GE, Howard BV, et al. Biomarker-calibrated dietary energy and protein intake association with diabetes risk among postmenopausal women from the Women's Health Initiative. Am J Clin Nutr. 2011; 94:1600–1606. [PubMed: 22071707]
19. Beasley J, LaCroix A, Neuhouser M, et al. Protein intake and incident frailty in the Women's Health Initiative Observational Study. J Am Geriatr Soc. 2010; 58:1063–1071. [PubMed: 20487071]
20. Beasley JM, Aragaki AK, LaCroix AZ, et al. Higher biomarker-calibrated protein intake is not associated with impaired renal function in postmenopausal women. J Nutr. 2011; 141:1502–1507. [PubMed: 21653574]
21. Prentice RL, Huang Y. Measurement error modeling and nutritional epidemiology association analyses. Canadian J Statistics. 2011; 39:498–509.

Table 1

Calibration equation coefficients (β), standard errors (SE), and percent of biomarker variation explained (R^2) from regression of log(biomarker) on log(self-report), body mass index, age, and ethnicity in the Women's Health Initiative Nutrition and Physical Activity Assessment Study among 450 postmenopausal women. Adapted from [5]. Log-energy (kcal/day) is centered by 7.27; log-protein (grams/day) is centered by 4.14; log-protein density (g/kcal) is centered by 2.85; BMI is centered by 28.2 kg/m^2; and age is centered by 70.9 years, in these analyses.

	Food Frequency				4DFR				24HRs(3)			
Variable	**β**	**SE**	**R^2**	**Adj R^2**	**β**	**SE**	**R^2**	**Adj R^2**	**β**	**SE**	**R^2**	**Adj R^2**
Energy (kilocalories)												
Intercept	7.614*	0.009			7.597*	0.009			7.607*	0.009		
FFQ	0.054*	0.017	3.8	6.5								
4DFR					0.161*	0.028	7.8	13.3				
24HR									0.101*	0.026	2.8	4.8
BMI	0.013*	0.001	26.9	45.9	0.013*	0.001	27.0	46.0	0.013*	0.001	28.7	48.9
Age	−0.010*	0.001	9.7	16.5	−0.009*	0.001	8.4	14.3	−0.009*	0.001	9.1	15.5
Black	−0.023	0.019			−0.024	0.018			−0.024	0.018		
Hispanic	−0.062*	0.021	1.3	2.2	−0.065*	0.020	1.5	2.6	−0.063*	0.020	1.5	2.6
Other minority	−0.041	0.040			−0.039	0.038			−0.038	0.039		
(Total)†			41.7	71.1			44.7	76.2			42.1	71.8
Protein (grams)												
Intercept	4.263	0.017			4.235	0.016			4.269	0.016		
FFQ	0.135*	0.021	8.4	16.4								
4DFR					0.465*	0.045	22.6	44.2				
24HR									0.404*	0.046	16.2	31.7
BMI	0.012*	0.002	5.8	11.4	0.012*	0.002	5.1	10.0	0.012*	0.002	5.8	11.4
Age	−0.012*	0.002	4.1	8.0	−0.009*	0.002	2.2	4.3	−0.011*	0.002	3.4	6.7
Black	−0.120*	0.038			−0.138*	0.034			−0.145*	0.035		
Hispanic	−0.078	0.040	2.0	3.9	−0.067	0.036	2.7	5.3	−0.069	0.037	3.0	5.9
Other minority	−0.018	0.076			0.012	0.070			−0.026	0.072		
(Total)†			20.3	39.7			32.7	63.8			28.4	55.6

	Food Frequency				4DFR				24HRs(3)			
Variable	**β**	**SE**	**R^2**	**Adj R^2**	**β**	**SE**	**R^2**	**Adj R^2**	**β**	**SE**	**R^2**	**Adj R^2**
Protein Density (% energy from protein)												
Intercept	2.652*	0.017			2.671*	0.017			2.687*	0.018		
FFQ	0.344*	0.068	6.5	38.5								
4DFR					0.488*	0.067	11.0	65.9				
24HR									0.393*	0.068	7.0	41.7
BMI	−0.002	0.002	0.6	3.3	−0.001	0.002	0.5	3.2	−0.001	0.002	0.4	2.3
Age	−0.002	0.002	0.04	0.2	−0.001	0.002	0.01	0.04	−0.003	0.002	0.1	0.7
Black	−0.100*	0.037			−0.130*	0.036			−0.133*	0.037		
Hispanic	−0.043	0.041	1.7	10.0	−0.035	0.040	2.8	17.0	−0.043	0.041	2.9	17.6
Other minority	−0.030	0.078			−0.006	0.075			−0.042	0.078		
(Total)†			8.7	52.1			14.4	86.1			10.4	62.3

BMI, body mass index; FFQ, food frequency questionnaire; 4DFR, four-day food record; 24HR, 24-hour dietary recall (3); UN, urinary nitrogen; Adj R^2, adjusted R^2 values (R^2 divided by log-biomarker correlation in reliability subsample).

*Coefficient differs from zero at p=0.05 significance level

†Total – percent of variation explained by all variables. R^2 values for specific variables arise from analyses with only these regression variables, with subsequent rescaling so that these R^2 values add to the total regression R^2. R^2 values for race/ethnicity pertain to comparisons among the four groups (white, black, Hispanic, other minority).

Table 2

Hazard ratio estimates for a 20% increment in calibrated FFQ energy, protein, and protein density in relation to coronary heart disease and stroke, based on data from 80,330 women enrolled in the WHI Dietary Modification Trial comparison group or the WHI Observational Study. Adapted With permission from: Prentice R.L. et al. Evaluation and Comparison of Food Records, Recalls, and Frequencies for Energy and Protein Assessment by Using Recovery Biomarkers. Am. J. Epidemiol.. 2011; [17] [21].

	Incident	Uncalibrated	Calibrated	Calibrated Consumption	Biomarker Measurement Error Correlation		
	Cases	Consumption	Consumption	Adjusted for BMI	0	0.375	0.50
Outcome		HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
				Energy			
CHD	1516	1.00	1.18	0.98	1.00	1.06	1.16
		(0.98, 1.03)	(1.04, 1.33)	(0.62, 1.27)	(0.81, 1.24)	(0.86, 1.26)	(0.94, 1.41)
Stroke	1224	0.99	0.86	0.61	0.71	0.76	0.85
		(0.97, 1.02)	(0.75, 1.00)	(0.41, 0.91)	(0.53, 0.95)	(0.60, 0.98)	(0.62, 1.17)
				Protein			
CHD	1516	0.99	1.01	0.89	0.93	0.99	
		(0.97, 1.01)	(0.92, 1.10)	(0.80, 0.99)	(0.84, 1.03)	(0.72, 1.35)	
Stroke	1224	0.99	0.89	0.87	0.88	0.89	
		(0.97, 1.01)	(0.82, 0.98)	(0.78, 0.98)	(0.79, 0.98)	(0.76, 1.05)	
				Protein Density			
CHD	1516	0.95	0.85	0.87	0.87	0.87	
		(0.91, 1.00)	(0.75, 0.97)	(0.78, 0.97)	(0.78, 0.98)	(0.77, 0.99)	
Stroke	1224	0.98	0.94	0.94	0.94	0.94	
		(0.93, 1.03)	(0.84, 1.06)	(0.83, 1.06)	(0.82, 1.07)	(0.82, 1.06)	

FFQ, food frequency questionnaire; WHI, Women's Health Initiative; CHD, coronary heart disease, comprised of nonfatal myocardial infarction plus coronary death; stroke, comprised of ischemic plus hemorrhagic stroke; HR, hazard ratio; CI, confidence interval