Practice of Epidemiology

Using Regression Calibration Equations That Combine Self-Reported Intake and Biomarker Measures to Obtain Unbiased Estimates and More Powerful Tests of Dietary Associations

Laurence S. Freedman*, Douglas Midthune, Raymond J. Carroll, Nataša Tasevska, Arthur Schatzkin†, Julie Mares, Lesley Tinker, Nancy Potischman, and Victor Kipnis

* Correspondence to Dr. Laurence S. Freedman, Biostatistics Unit, Gertner Institute for Epidemiology and Health Policy Research, Tel Hashomer 52621, Israel (e-mail: lsf@actcom.co.il).
†Deceased.

Initially submitted February 20, 2011; accepted for publication June 24, 2011.

The authors describe a statistical method of combining self-reports and biomarkers that, with adequate control for confounding, will provide nearly unbiased estimates of diet-disease associations and a valid test of the null hypothesis of no association. The method is based on regression calibration. In cases in which the diet-disease association is mediated by the biomarker, the association needs to be estimated as the total dietary effect in a mediation model. However, the hypothesis of no association is best tested through a marginal model that includes as the exposure the regression calibration-estimated intake but not the biomarker. The authors illustrate the method with data from the Carotenoids and Age-Related Eye Disease Study (2001–2004) and show that inclusion of the biomarker in the regression calibration-estimated intake increases the statistical power. This development sheds light on previous analyses of diet-disease associations reported in the literature.

bias (epidemiology); carotenoids; cataract; lutein; measurement error; sample size

Abbreviations: BMI, body mass index; CAREDS, Carotenoids and Age-Related Eye Disease Study; FFQ, food frequency questionnaire; RC, regression calibration; WHI, Women’s Health Initiative.

Dietary measurement error causes serious challenges to the detection of associations between diet and disease in epidemiologic studies. Estimated relative risks are attenuated and statistical power is reduced (1). Moreover, increasing the sample size provides only a partial remedy, as the attenuated estimates of relative risk are often as low as 1.10–1.25 and, even when statistically significant, might be indistinguishable from the effects of unknown confounders (2). Methods to reduce error in dietary measurements are therefore of primary importance.

To that aim, in previous work (3, 4), we proposed using Howe’s method or principal components for combining dietary reports with dietary biomarkers. In computer simulations and in a real example from the Carotenoids and Age-Related Eye Disease Study (CAREDS), we demonstrated substantial increases in statistical power with this method. However, the method is limited in 3 ways. First, the relative risks derived cannot be translated directly into relative risks per added unit of intake. Second, the method does not always increase power but rather sometimes decreases it. Third, the method does not derive from the usual statistical framework for modeling measurement error.

In the present article, we propose an alternative method for combining dietary self-reports and biomarkers derived from regression calibration (5), a well-known statistical approach to solving measurement error problems. Through theory and computer simulations, we show that, given adequate control for confounding, this approach gives nearly unbiased relative risk estimates and provides a significance test with power equal to or greater than that of tests based on dietary self-report alone. We apply the method to study the negative association of lutein plus zeaxanthin intake with nuclear cataracts that we reported in a previous article (4).

We also compare our method with a closely related proposal of Prentice et al. (6) and clarify methodological issues.
related to including a biomarker in the regression calibration equation when that same biomarker is involved in the disease model.

**MATERIALS AND METHODS**

**Study population**

CAREDS is an ancillary study of the Women’s Health Initiative (WHI) Observational Study (7, 8). The CAREDS population included women enrolled at 3 sites: University of Wisconsin (Madison, Wisconsin), University of Iowa (Iowa City, Iowa), and Kaiser Center for Health Research (Portland, Oregon). Of the 3,143 eligible women, 2,005 agreed to participate. Full details of the study design have been published previously (9). All procedures conformed to the Declaration of Helsinki and were approved by institutional review boards at each institution.

**Assessments**

Dietary intake was assessed at the WHI Observational Study baseline (1994–1998) by using the WHI semiquantitative food frequency questionnaire (FFQ), which had been pretested (10). Serum samples were collected after subjects had fasted for 10 or more hours at the WHI baseline examinations (1994–1998) and were analyzed for lutein and zeaxanthin (sum of their trans isomers) (9). Serum lutein and zeaxanthin measurements were available for 1,787 women. These women comprised the data set for this study. Participants underwent lens photography and eye examinations during the CAREDS baseline study visits between 2001 and 2004 (9) and completed a questionnaire that included questions about time of cataract surgery in each eye, physician-diagnosed history of cataracts, and personal characteristics. The primary outcome was nuclear cataract, defined as a nuclear sclerosis severity score of 4 or greater in the worst eye or a history of cataract extraction in either eye.

**Confounders**

Potential confounders used in the CAREDS analyses relating nuclear cataracts to lutein/zeaxanthin were age, smoking, iris color, body mass index (BMI, measured as weight in kilograms divided by height in meters squared), multivitamin use, physical activity level, hormone replacement therapy, and pulse pressure (9). In the present analysis, we adjust for the 2 strongest confounders, age and smoking. We considered adding BMI to the analysis, but this did not materially change the results.

**Statistical methods**

**Measurement error model.** We assume there are both a single usual dietary intake of interest and a single biomarker measure related to the intake. The true values of these are $X_I$ and $X_M$, respectively ($I$ for intake and $M$ for marker). Each is measured with error, the measurements being denoted by $W_I$ and $W_M$, respectively. In our example, these measurements are log FFQ-reported lutein plus zeaxanthin intake and log serum level of lutein plus zeaxanthin.

To deal with the effects of the measurement error, we have to know the measurement error model. We assume it takes the following form:

$$W_I = \alpha_{10} + \alpha_{11}X_I + \alpha_{12}X_M + \epsilon_1,$$

$$W_M = \alpha_{20} + \alpha_{21}X_I + \alpha_{22}X_M + \epsilon_2.$$  \hspace{1cm} (1)

This model is quite general, allowing the intake measurement error to depend on true intake ($\alpha_{11} \neq 1$) and even on true biomarker level ($\alpha_{12} \neq 0$); similarly, the biomarker measurement error may depend on true biomarker level ($\alpha_{22} \neq 1$) and even on true intake level ($\alpha_{21} \neq 0$). The random error ($\epsilon_1$ and $\epsilon_2$) terms (with zero means) are assumed to be independent of each other, true intake, true biomarker level, and disease outcome, so that the measurement error is non-differential. The random variables ($X_I, X_M, W_I,$ and $W_M$) are assumed to have a multivariate normal distribution with a known covariance matrix. In practice, these variances and covariances, as well as the $\alpha$-parameters in model 1, can be estimated by using data from feeding studies in conjunction with the data on $W_I$ and $W_M$ from the main study. An example of the derivations of these values from the literature is given for lutein plus zeaxanthin in the Appendix to Freedman et al. (3).

In our CAREDS example, the measurement error model was simpler, with $\alpha_{12} = \alpha_{21} = 0$, $\alpha_{20} = 0$, and $\alpha_{22} = 1$. Each assumption seemed appropriate in this context. Setting $\alpha_{12} = 0$ corresponded to assuming that error in dietary reporting of lutein/zeaxanthin intake was unrelated to the serum level, and setting $\alpha_{21} = 0$ corresponded to assuming that error in measured serum level was unrelated to the intake level. Setting $\alpha_{20} = 0$ and $\alpha_{22} = 1$ corresponded to assuming that the measured serum level was an unbiased measure of true serum level.

**Disease model.** Denote the disease outcome variable by $Y$. We assume that disease is related to the exposure of interest $X_I$ through a generalized linear regression model:

$$h[E(Y|X_I, X_M, Z)] = \beta_0 + \beta_1X_I + \beta_2X_M + \beta_ZZ,$$  \hspace{1cm} (2)

where $E$ denotes expectation, $h$ denotes the link function (e.g., logistic for binary variables or the identity for a continuous variable), and $Z$ represents one or more confounders, measured exactly.

We specifically include the biomarker variable $X_M$ in disease model 2 because we assume that the biomarker mediates, at least partially, the effect of dietary intake on disease. See Figure 1 for the causal path diagram. This assumption

![Figure 1. Causal pathway diagram describing the relations among dietary intake, biomarker level, confounders, and disease.](https://academic.oup.com/aje/article-abstract/174/11/1238/110812)
often seems biologically plausible; for example, if the biomarker is a serum level of the nutrient of interest, then the effect of the nutrient intake on disease will likely be at least partially mediated through the biomarker. We assume that, if the biomarker level is influenced by other factors associated with disease, those factors are included in the covariates \( Z \) (Figure 1); in other words, we assume that we can control for confounding.

Under this causal model (Figure 1), the quantity of most interest is the coefficient for the total association of dietary intake with disease outcome, given by \( \beta_1 \) in the following model:

\[
h[E(Y|X_I, Z)] = \beta_0 + \beta_1^* X_I + \beta_2 Z.
\]

(3)

It can be shown that both model 3 and the following equation for \( \beta_1 \) hold exactly when the disease model is a linear regression and approximately when the disease model is nonlinear, such as for logistic and Cox regression:

\[
\beta_1^* = \beta_1 + \gamma \beta_2,
\]

(4)

where \( \gamma \) is the coefficient of \( X_I \) in the linear regression of \( X_M \) on \( X_I \) and \( Z \).

In the absence of measurement error, we could estimate \( \beta_1^* \) simply by omitting \( X_M \) from model 2 and estimating the coefficient for \( X_I \) in model 3. However, as detailed in the next subsection, if we utilize this same strategy using measurement error adjustment, then we can get a biased estimate of \( \beta_1^* \).

Estimating the total association of dietary intake with disease. Regression calibration (RC) (5) is now widely used to adjust estimates of regression coefficients for measurement error in explanatory variables. Suppose we are interested in estimating \( \beta_1^* \) in model 3. The central idea is to use as the explanatory variable in the regression not \( W_I \) but rather the expectation of the true value \( X_I \) conditional on its measurement \( W_I \), as well as the covariates \( Z \), written as \( E(X_I|W_I, Z) \). This is essentially a prediction of the unknown \( X_I \) using its mismeasured value together with the confounders. As long as the measurement error is nondifferential, the method yields regression estimates that in large samples are unbiased (for linear disease models) or nearly unbiased (for nonlinear disease models and moderate-sized relative risks) (5). However, the statistical power for detecting an association is not increased and can actually decrease slightly (5).

Kipnis et al. (11) demonstrated that RC can be extended by including other variables in the “prediction” of \( X_I \). Such extra variables increase the precision with which \( X_I \) can be predicted and thereby also increase the precision with which \( \beta_1^* \) is estimated. One caveat, however, is that the extra predictors should provide no further information about the disease outcome over and above that provided by \( X_I \) itself and the confounders \( Z \). Without this condition, the method yields a biased estimate of \( \beta_1^* \), and the only remedy is to include those extra predictors in the disease model as well.

In the present article, we explore the use of the biomarker \( W_M \) as an extra predictor of \( X_I \). We compare several methods of estimating \( \beta_1^* \):

1. Unadjusted for measurement error, using the estimated coefficient for \( W_I \) in the model \( h[E(Y)] = \beta_{0W} + \beta_{1W} W_I + \beta_{2W} Z \);
2. Usual RC, using the estimated coefficient for \( E(X_I|W_I, Z) \) in the model \( h[E(Y)] = \beta_{0R} + \beta_{1R} E(X_I|W_I, Z) + \beta_{2R} Z \);
3. Enhanced RC, with the biomarker used for prediction, using the estimated coefficient for \( E(X_I|W_I, W_M, Z) \) in the model \( h[E(Y)] = \beta_{0E} + \beta_{1E} E(X_I|W_I, W_M, Z) + \beta_{2E} Z \) (the method used by Prentice et al. (6) in an example where the dietary intake was total energy and the “biomarker” was BMI); and
4. A newly proposed method, in which enhanced RC is used to estimate \( \beta_1 \) and \( \beta_2 \) of model 2 through the model \( h[E(Y)] = \beta_{0N} + \beta_{1N} E(X_I|W_I, W_M, Z) + \beta_{2N} E(X_M|W_I, W_M, Z) + \beta_{2Z} Z \) (\( \beta_1 \) is then estimated by \( \hat{\beta}_{1N} + \gamma \hat{\beta}_{2N} \), where the hat denotes the estimated value. Note that in the regression model, we use \( E(X_M|W_I, W_M, Z) \) in place of \( W_M \) to account for any measurement error in the biomarker.

In the Web Appendix, Part A (available at http://aje.oxfordjournals.org/), we present the approximate expected values of these 4 estimators. This enables us to predict the following regarding the bias in estimating \( \beta_1^* \):

1. The unadjusted estimate, \( \hat{\beta}_{1U} \), is nearly unbiased only if there is no measurement error in dietary intake.
2. The usual RC estimate, \( \hat{\beta}_{1R} \), is nearly unbiased if either \( \beta_2 \) is zero (i.e., if there is no mediation by the biomarker) or \( x_{12} \) is zero (i.e., if the biomarker provides no information about reported dietary intake over and above that provided by the true intake). The latter scenario is sometimes plausible but not in the example of energy intake and BMI (6).
3. The enhanced RC estimate, \( \hat{\beta}_{1E} \), is nearly unbiased if \( \beta_2 \) is zero (i.e., if there is no mediation by the biomarker) but not in any other plausible scenarios.
4. The newly proposed estimate is nearly unbiased under the more general conditions of the measurement error model 1 and the disease model 2.

In the Results, we present the results of computer simulations that verify these predictions.

To apply these methods to the CAREDs example, we must develop the prediction models, that is, the quantities \( E(X_I|W_I, Z) \), \( E(X_I|W_I, W_M, Z) \), and \( E(X_M|W_I, W_M, Z) \) involved in usual RC, enhanced RC, and the newly proposed method. See the Web Appendix, Part B.

Testing the null hypothesis of no diet-disease association. Each of the 4 estimation methods described above can also be used to test the hypothesis of no diet-disease association. In each case, the test is obtained by comparing the ratio of the estimate to its standard error with the standard normal distribution. The estimate’s standard error may be computed by using bootstrap methods or, when the measurement error model parameters are assumed known, from the usual model-based estimates.

Two questions arise: 1) Which of the 4 tests are valid? and 2) Which among the valid tests is the most statistically powerful? Answering these questions requires careful definition of the null hypothesis. Specifically, we mean that not only is \( \beta_1^* \) zero, but also that its 2 components \( \beta_1 \) and \( \beta_2 \) in model 2 are both zero. It is theoretically possible that \( \beta_1^* \), equals zero even when these 2 components are not zero, namely, when \( \beta_1 = -\gamma \beta_2 \). However, it would be highly
unusual for the direct effect of diet on disease (the part not mediated through the biomarker) to be in the opposite direction of the indirect effect (the part mediated) with the two effects in precisely the appropriate ratio to cancel each other. Thus, we concentrate on the more plausible hypothesis that the biomarker mediates through the biomarker) to be in the opposite direction of the indirect effect (the part not mediated by the biomarker).

### RESULTS

#### CAREDS example

**Regression calibration equations.** Equations for predicting true intake (and true biomarker level) are required to implement usual RC, enhanced RC, and the newly proposed method. For usual RC, the quantity $E(X_t | W_t, Z)$ is required, where $W_t$ is the log FFQ report of lutein/zeaxanthin intake and $Z$ represents the covariates age ($y$) and smoking ($0 =$ nonsmoker, $1 =$ former smoker, $2 =$ current smoker). We obtained the equation:

$$\text{Predicted intake} = 0.355 W_t + 0.00560 age - 0.101 smoking.$$  

For enhanced RC, the biomarker level was included in the prediction, and we obtained the equation:

$$\text{Predicted intake} = 0.242 W_t + 0.515 W_M + 0.00692 age - 0.0954 smoking.$$  

Note that the coefficient for $W_M$, the log measured serum level, is relatively large, showing its major role in the prediction of true intake.

For the newly proposed method, we also needed a prediction equation for true serum lutein/zeaxanthin, obtained as follows:

$$\text{Predicted serum level} = 0.0508 W_t + 0.769 W_M - 0.000592 age - 0.00234 smoking.$$  

Note that the coefficient of $W_t$, the log FFQ-reported intake, is small, showing its minor role in predicting serum level.

**Estimates of risk and significance tests.** Estimates of the odds ratio of nuclear cataract associated with a doubling of lutein/zeaxanthin intake derived from the 4 methods are shown in Table 2. The unadjusted estimate of 0.89 was closer to the null value of 1.0 than were the estimates of the other methods, and it just achieved statistical significance ($P = 0.038$). Usual RC yielded a stronger (negative) association (odds ratio = 0.72) but the same level of significance ($P = 0.038$). Enhanced RC estimated an even stronger association (odds ratio = 0.70) that was highly significant ($P = 0.002$). The newly proposed method estimated a weaker association than did the enhanced

### Table 1. Parameters for Simulation Study of Both Dietary and Serum Lutein/Zeaxanthin Based on the Models 1 and 2a, Carotenoids in Age-Related Eye Disease Study, 2001–2004

<table>
<thead>
<tr>
<th>Set</th>
<th>$\alpha_{11}$</th>
<th>$\alpha_{12}$</th>
<th>$\alpha_{21}$</th>
<th>$\alpha_{22}$</th>
<th>Variance of $\epsilon_1$</th>
<th>Variance of $\epsilon_2$</th>
<th>Variance of $X_t$</th>
<th>Variance of $X_M$</th>
<th>$\gamma$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No mediation</td>
<td>0.71</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.36</td>
<td>0.05</td>
<td>0.25</td>
<td>0.19</td>
<td>0.60</td>
<td>-1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Full mediation</td>
<td>0.71</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.36</td>
<td>0.05</td>
<td>0.25</td>
<td>0.19</td>
<td>0.60</td>
<td>0.00</td>
<td>-1.20</td>
</tr>
<tr>
<td>Partial mediation</td>
<td>0.71</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.36</td>
<td>0.05</td>
<td>0.25</td>
<td>0.19</td>
<td>0.60</td>
<td>-0.48</td>
<td>-0.63</td>
</tr>
<tr>
<td>No effect</td>
<td>0.71</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.36</td>
<td>0.05</td>
<td>0.25</td>
<td>0.19</td>
<td>0.60</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*a Measurement error model (model 1): $W_t = \alpha_{11} X_1 + \alpha_{12} X_2 + \epsilon_1$, $W_M = \alpha_{21} X_1 + \alpha_{22} X_2 + \epsilon_2$. Disease model (model 2): $h[E(Y, X_t, X_M)] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 Z$, where $X_t$ and $X_M$ are the true values of intake and biomarker, respectively; $W_t$ and $W_M$ are their measured values; $Y$ is the disease indicator; the $\epsilon$'s are random errors; and $\gamma$ is the coefficient of $X_t$ in the linear regression of $X_M$ on $X_t$ and $Z$. 

Using Calibration to Combine Diet Reports and Markers. The simulation was designed to mimic data from CAREDS. Parameters for the measurement error and disease models, derived from the literature (3), are shown in Table 1. The different combinations of ($\beta_1$, $\beta_2$) values were designed to represent 4 scenarios for the effect of intake on disease: 1) not mediated by the biomarker, 2) fully mediated by the biomarker, 3) partially mediated by the biomarker, and 4) zero. The RC models were assumed known and were calculated from the measurement error parameters.

For each simulation, a study with 500 individuals and measurements ($Y, W_t, W_M$) was generated from the models. Then, the 4 estimation methods were applied to the data, leading to 4 estimates of $\beta_1$. This was repeated 1,000 times, and the means and standard deviations of each of the 4 estimators over the 1,000 repetitions were calculated. To estimate statistical power, Wald chi-squared statistics were computed, and the proportion of simulations in which the statistic exceeded the 95th percentile was calculated.
RC method (odds ratio = 0.74), with a level of significance similar to that of usual RC ($P = 0.046$).

Which of the 4 methods should one choose? For estimation, the newly proposed method is the only one that is nearly unbiased under general models 1 and 2, but in our particular case in which $\alpha_{21} = 0$, the usual RC method is also nearly unbiased. Therefore, one may choose between them, and, because its standard error is smaller, the newly proposed method would be preferable. With regard to significance testing, all the methods are nearly unbiased, and the natural choice is the enhanced RC method because it is the most powerful.

**Sample size savings.** We calculated the predicted ratio of sample size required using enhanced RC to that required using usual RC,

$$\frac{\text{var}[E[X_j|W_j,Z]]}{\text{var}[E[X_j|W_j, W_M, Z]]} \approx \frac{0.0683}{0.1250} = 0.55.$$  

This value was calculated without reference to the outcome variable. However, CAREDS data allowed us to calculate the sample size savings in relation to testing the association with nuclear cataracts. The sample size ratio for the enhanced RC method versus the usual RC method was estimated as 0.43 (Table 2), which was not dissimilar to the predicted value of 0.55.

**Computer simulations**

Results regarding bias in the estimated risk parameter resembled those predicted by theory (Table 3). The unadjusted estimate was attenuated, except under zero dietary effect. The usual RC method gave nearly unbiased estimates in all scenarios because we set the measurement error parameter $\alpha_{21}$ at zero. The enhanced RC method gave nearly unbiased estimates under no mediation of the dietary effect through the biomarker and also under zero dietary effect, but when mediation occurred, the estimate was biased and inflated away from the null. The newly proposed method gave nearly unbiased estimates in all scenarios.

The precisions of the estimates differed markedly. The standard deviations of the usual RC and newly proposed estimates were similar and larger than those of enhanced RC estimates. Each of the 4 significance tests yielded approximately 5% of significant results under the null hypothesis (Table 4). The tests based on the unadjusted method and RC method were identical and had lower statistical power than did the test based on the enhanced RC method (Table 4). The test based on the newly proposed method had slightly higher power than did the RC method.

**DISCUSSION**

We have described a method of combining self-reports and biomarkers, based on RC that, under reasonable assumptions, provides 1) a nearly valid significance test of the diet-disease association with increased power and 2) nearly unbiased estimates of relative risks or odds ratios for the association.

The method relies on prior knowledge or estimation of the statistical model describing the measurement error in self-reports of the dietary intake and in biomarker levels related to dietary intake. For the few existing recovery biomarkers

### Table 2. Logistic Regression Analyses Relating Nuclear Cataracts to Dietary Lutein/Zeaxanthin Intake in the Carotenoids in Age-Related Eye Disease Study, 2001–2004

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>Coefficient (SE)</th>
<th>Odds Ratio*</th>
<th>Z Value</th>
<th>Sample-Size Ratio(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>-0.165 (0.080)</td>
<td>0.89</td>
<td>-2.07</td>
<td>1.0</td>
</tr>
<tr>
<td>Regression calibration</td>
<td>-0.464 (0.225)</td>
<td>0.72</td>
<td>-2.07</td>
<td>1.0</td>
</tr>
<tr>
<td>Enhanced regression calibration</td>
<td>-0.506 (0.161)</td>
<td>0.70</td>
<td>-3.15</td>
<td>0.43</td>
</tr>
<tr>
<td>Newly proposed</td>
<td>-0.436 (0.218)</td>
<td>0.74</td>
<td>-2.00</td>
<td>1.07</td>
</tr>
</tbody>
</table>

* Odds ratio of disease for a doubling of the dietary intake of lutein/zeaxanthin.

\(^b\) The ratios of sample sizes required to give the same statistical power as that based on the unadjusted analysis.

### Table 3. Simulated Means and Standard Deviations of Estimates of the Marginal Effect $\beta_1$ of Dietary and Serum Lutein/Zeaxanthin on Eye Disease for 4 Different Methods of Estimation\(^a\)

<table>
<thead>
<tr>
<th>Set</th>
<th>True $\beta_1$</th>
<th>Unadjusted</th>
<th>Regression Calibration</th>
<th>Enhanced Regression Calibration</th>
<th>Newly Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (SE)</td>
<td>SD</td>
<td>Mean (SE)</td>
<td>Mean (SE)</td>
</tr>
<tr>
<td>No mediation</td>
<td>-1.00</td>
<td>-0.349 (0.004)</td>
<td>0.135</td>
<td>-0.951 (0.008)</td>
<td>0.259</td>
</tr>
<tr>
<td>Full mediation</td>
<td>-0.72</td>
<td>-0.257 (0.004)</td>
<td>0.132</td>
<td>-0.688 (0.008)</td>
<td>0.254</td>
</tr>
<tr>
<td>Partial mediation</td>
<td>-0.86</td>
<td>-0.303 (0.004)</td>
<td>0.131</td>
<td>-0.826 (0.008)</td>
<td>0.258</td>
</tr>
<tr>
<td>No effect</td>
<td>0.00</td>
<td>-0.006 (0.004)</td>
<td>0.128</td>
<td>-0.017 (0.011)</td>
<td>0.351</td>
</tr>
</tbody>
</table>

Abbreviations: SD, standard deviation; SE, standard error.

\(^a\) There were 1,000 simulated studies per set with 500 individuals per study.
(the doubly labeled water technique (12) for measuring energy intake and 24-hour urinary nitrogen (13) and potassium (14) for measuring protein and potassium intake), this method could be easily applied given the biomarkers’ known quantitative relation to intake (13), although the cost or effort to perform these tests in very large numbers may be prohibitive. For the newly developed predictive biomarker for sugars (15), estimation of the measurement error parameters has been recently described (16). In our example, we extracted such prior knowledge from the literature on carotenoid feeding studies, validation studies of dietary reporting of carotenoids, and cohort studies that investigated carotenoid-disease associations (3). For other concentration biomarkers, such as other serum carotenoids or vitamin C or adipose tissue fatty acids, a similar exercise using previous feeding studies could be attempted; otherwise, new feeding studies will be needed to develop the RC equations lying at the heart of the method. One such feeding study is now being conducted (17). The method is not applicable to foods or food patterns that have no known specific biomarkers.

When the parameters of the measurement error model are estimated from a feeding study, the limited size of the study often limits the precision of the estimates. This uncertainty transfers to the risk estimates obtained from the RC adjustment. One should then use a method to adjust the standard error of the risk estimate to include this extra uncertainty, such as the bootstrap or stacking equations (see Appendix B.3 of Carroll et al. (5)).

The method we propose is linked to a previous proposal to use principal components or Howe’s method to combine different error-prone measures of dietary intake. Our results, which showed that enhanced RC can yield reductions in sample size of approximately 50%, are similar to the savings found using Howe’s method applied to the same data set (4). However, this similarity is serendipitous. It happened that in this data set, Howe’s method yielded a dietary index close to the RC prediction of dietary intake based on self-reported intake and serum level, and consequently provided a nearly optimal analysis. This will not always happen, and neither principal components nor Howe’s method is guaranteed to increase statistical power. In many cases, both methods can actually decrease it. In contrast, enhanced RC is expected always to increase power provided the biomarker improves prediction of dietary intake.

Although the enhanced RC method yields a valid and more powerful significance test of the diet-disease association, it does not in general provide an unbiased estimate of the risk parameter. Whenever there is mediation of the dietary effect through the biomarker, which is often expected, the estimate becomes inflated. We have provided a new method that yields an unbiased estimate, albeit with lower precision than that provided by the enhanced RC method. When the biomarker is uncorrelated with dietary reporting error, the usual RC estimate (based on self-report alone) will also be unbiased. We recommend reporting one of these unbiased estimates together with the result of the significance test based on enhanced RC.

Prentice et al. (6) presented an analysis of the association between energy intake and several invasive cancers based on a calibration equation for energy intake that includes BMI. The authors discussed whether BMI was a confounder or a mediator of the diet-disease association. Assuming it was a mediator, they estimated hazard ratios by using a method that corresponded to enhanced RC. According to the results of the present study, the significance tests of the diet-disease association were valid, but the estimated hazard ratios were biased. If BMI were a confounder, then the risk quantity of interest would no longer be \( \beta_1 \) but \( \beta_1 + \gamma \beta_2 \) in model 2, the confounder-adjusted dietary risk parameter.

Possible confounding of the diet-disease association through the biomarker is the most serious obstacle to using our approach. If there are risk factors for disease that also affect the biomarker, then introducing the biomarker into the prediction of dietary intake while not controlling for these risk factors in the disease model will lead to biased estimation with unknown direction of the bias. Biomarkers known to bear a strong relation to dietary intake, such as recovery (18) and predictive (15) biomarkers, will be largely immune from such concern, but concentration biomarkers are affected by complex metabolic pathways in their regulation and will always be subject to concerns about confounding. If a strong risk factor for the disease is known to affect the biomarker, that risk factor must be included in the disease risk model so as to avoid ascribing its effect to nutritional components. In the
CAREDS example, smoking was included in the model because it is associated with nuclear cataract and might also lead to depletion of lutein and zeaxanthin in blood, as it is a source of free radicals and oxidative stress (19).

Another challenge is the need to specify a measurement error model, such as model 1. Naturally, such a model could be incomplete and might omit influential explanatory variables. For a discussion of this challenge and related cost issues, see the Web Appendix, Part C.

In summary, a major obstacle facing the field of nutritional epidemiology is the loss of statistical power for detecting diet-disease associations that result from measurement error. With careful use, the methods described in the present article could yield more powerful tests of such associations together with reliable risk estimates. Use of these methods in other branches of epidemiology is discussed briefly in the Web Appendix, Part C.

ACKNOWLEDGMENTS

Author affiliations: Biostatistics Unit, Gertner Institute for Epidemiology and Health Policy Research, Tel Hashomer, Israel (Laurence S. Freedman); Division of Cancer Prevention, National Cancer Institute, Bethesda, Maryland (Douglas Midthune, Victor Kipnis); Department of Statistics, Texas A&M University, College Station, Texas (Raymond J. Carroll); Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland (Nataša Tasevská, Nancy Potischman); Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland (Arthur Schatzkin); Department of Ophthalmology and Visual Sciences, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, Wisconsin (Julie Mares); and Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington (Lesley Tinker).

This work was supported by the National Institutes of Health (under contract HHSN261200633000 to L. S. F.), the National Cancer Institute (grant R27-CA057030 to R. J. C.), the National Eye Institute (grants EY013018 and EY016886), the National Heart, Lung, and Blood Institute (for support of the Women's Health Initiative), and by Research to Prevent Blindness.

Women's Health Initiative Investigators—Program Office (National Heart, Lung, and Blood Institute, Bethesda, Maryland): Elizabeth Nabel, Jacques Rossouw, Shari Ludlam, Joan McGowan, Leslie Ford, and Nancy Geller. Clinical Coordinating Centers—Fred Hutchinson Cancer Research Center, Seattle, Washington: Ross Prentice, Garnet Anderson, Andrea LaCroix, Charles L. Kooperberg, Ruth E. Patterson, and Anne McTiernan; Medical Research Labs, Highland Heights, Kentucky: Evan Stein; and University of California at San Francisco, San Francisco, California: Steven Cummings. Clinical Centers—Albert Einstein College of Medicine, Bronx, New York: Sylvia Wassertheil-Smoller; Baylor College of Medicine, Houston, Texas: Aleksandar Rajkovic; Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts: JoAnn E. Manson; Brown University, Providence, Rhode Island: Charles B. Eaton; Emory University, Atlanta, Georgia: Lawrence Phillips; Fred Hutchinson Cancer Research Center, Seattle, Washington: Shirley Beresford; George Washington University Medical Center, Washington, DC: Lisa Martin; Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California: Rowan Chlebowski; Kaiser Permanente Center for Health Research, Portland, Oregon: Yvonne Michael; Kaiser Permanente Division of Research, Oakland, California: Bette Caan; Medical College of Wisconsin, Milwaukee, Wisconsin: Jane Morley Kotchen; MedStar Research Institute/Howard University, Washington, DC: Barbara V. Howard; Northwestern University, Evanston, Illinois: Linda Van Horn; Rush Medical Center, Chicago, Illinois: Henry Black; Stanford Prevention Research Center, Stanford, California: Marcia L. Stefanick; State University of New York at Stony Brook, Stony Brook, New York: Dorothy Lane; Ohio State University, Columbus, Ohio: Rebecca Jackson; University of Alabama at Birmingham, Birmingham, Alabama: Cora E. Lewis; University of Arizona, Phoenix, Arizona: Cynthia A. Thomson; University at Buffalo, Buffalo, New York: Jean Wactawski-Wende; University of California at Davis, Sacramento, California: John Robbins; University of California at Irvine, California: F. Allan Hubbell; University of California at Los Angeles, Los Angeles, California: Lauren Nathan; University of California at San Diego, La Jolla/Chula Vista, California: Robert D. Langer; University of Cincinnati, Cincinnati, Ohio: Margery Gass; University of Florida, Gainesville/Jacksonville, Florida: Marian Limerich; University of Hawaii, Honolulu, Hawaii: J. David Curb; University of Iowa, Iowa City, Iowa: Robert Wallace; University of Massachusetts/Fallon Clinic, Worcester, Massachusetts: Judith Ockene; University of Medicine and Dentistry of New Jersey, Newark, New Jersey: Norman Lasser; University of Miami, Miami, Florida: Mary Jo O'Sullivan; University of Minnesota, Minneapolis, Minnesota: Karen Margolis; University of Nevada, Reno, Nevada: Robert Brunner; University of North Carolina, Chapel Hill, North Carolina: Gerardo Heiss; University of Pittsburgh, Pittsburgh, Pennsylvania: Lewis Kuller; University of Tennessee Health Science Center, Memphis, Tennessee: Karen C. Johnson; University of Texas Health Science Center, San Antonio, Texas: Robert Brzyski; University of Wisconsin, Madison, Wisconsin: Glorina E. Sarto; Wake Forest University School of Medicine, Winston-Salem, North Carolina: Mara Vitolins; and Wayne State University School of Medicine/Hutzel Hospital, Detroit, Michigan: Michael Simon. Women's Health Initiative Memory Study (Wake Forest University School of Medicine, Winston-Salem, North Carolina): Sally Shumaker.

Conflict of interest: none declared.

REFERENCES


