HHS Public Access

Author manuscript

Biometrics. Author manuscript; available in PMC 2016 December 27.

Published in final edited form as:

Biometrics. 2016 December ; 72(4): 1369–1377. doi:10.1111/biom.12524.

Moment Reconstruction and Moment-Adjusted Imputation When Exposure is Generated by a Complex, Nonlinear Random Effects Modeling Process

Cornelis J. Potgieter[1,*], **Rubin Wei**[2,**], **Victor Kipnis**[3,***], **Laurence S. Freedman**[4,****], and **Raymond J. Carroll**[5,*****]

[1]Department of Statistical Science, Southern Methodist University, Dallas TX 75275

[2]Eli Lilly and Company, Indianapolis IN 46285

[3]Biometry Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, MD 20814

[4]Gertner Institute for Epidemiology and Health Policy Research, Tel Hashomer, Israel

[5]Department of Statistics, Texas A&M University, College Station, TX 77843, and School of Mathematical and Physical Sciences, University of Technology, Sydney, NSW, 2007

Summary

For the classical, homoscedastic measurement error model, moment reconstruction (Freedman et al., 2004, 2008) and moment-adjusted imputation (Thomas et al., 2011) are appealing, computationally simple imputation-like methods for general model fitting. Like classical regression calibration, the idea is to replace the unobserved variable subject to measurement error with a proxy that can be used in a variety of analyses. Moment reconstruction and moment-adjusted imputation differ from regression calibration in that they attempt to match multiple features of the latent variable, and also to match some of the latent variable's relationships with the response and additional covariates. In this note, we consider a problem where true exposure is generated by a complex, nonlinear random effects modeling process, and develop analogues of moment reconstruction and moment-adjusted imputation for this case. This general model

*cpotgieter@smu.edu
**weirubin@gmail.com
***kipnisv@mail.nih.gov
****lsf@actcom.co.il
*****carroll@stat.tamu.edu

Supplementary Material

Web Appendices, and Figures referenced in Sections 1, 3.5, 4.1, 4.2 and 6 are available with this paper at the Biometrics website on Wiley Online Library. The NIH-AARP Diet and Health Study data are available via a proposal process, see http://dietandhealth.cancer.gov. Fitting the method of Zhang et al. (2011) to the NIH-AARP calibration data can be done either by a SAS macro available at http://epi.grants.cancer.gov/diet/usualintakes (click on the "Read more about the NCI method"), or by Matlab code available at the Biometrics website as a zip file. The zip file has three Word files: (a) Read_Me_Extracting_Data_Sets.doc, which gives instruction on how to extract the relevant data from the NIH-AARP Study, once it is obtained; (b) Read_Me_Calibration_Study.doc, which gives instructions for what to do with the calibration study data; and (c) Read_Me_Primary_Study.doc, which lists the programs that are used and gives their order. However, the data are being regularly updated, with many more colorectal cancer cases being found, and so results applied to the most recent data will not match the results we found with much older data. **Supplementary Material** includes technical details of fitting the logistic regression model, figures of Bspline fits, and a simulation where it is shown that two forms of regression calibration which uses dummy variables can have substantial bias.

includes classical measurement errors, Berkson measurement errors, mixtures of Berkson and classical errors and problems that are not measurement error problems, but also cases where the data generating process for true exposure is a complex, nonlinear random effects modeling process. The methods are illustrated using the National Institutes of Health-AARP Diet and Health Study where the latent variable is a dietary pattern score called the Healthy Eating Index - 2005. We also show how our general model includes methods used in radiation epidemiology as a special case. Simulations are used to illustrate the methods.

Keywords

Berkson-type error; Classical measurement error; Computer models; Healthy Eating Index-2005; Latent variable models; Moment-adjusted imputation; Moment reconstruction; Nutritional Epidemiology

1. Introduction

In measurement error modeling, moment reconstruction (Freedman et al., 2004, 2008) and moment-adjusted imputation (Thomas et al., 2011, 2013) are appealing, computationally simple imputation-like methods for general model fitting. Let the outcome be Y, the latent variables of interest be the vector $\mathbf{X}$, covariates measured without error $\mathbf{Z}$, and let $\mathbf{Q}$ be the mismeasured version of $\mathbf{X}$. Both methods assume the classical additive error model that $\mathbf{Q} = \mathbf{X} + \mathbf{U}$, where $\mathbf{U}$ is independent of (Y, $\mathbf{X}$, $\mathbf{Z}$) and is thus homoscedastic, although Freedman et al. (2008) extend moment reconstruction when $\mathbf{Q}$ has a linear bias that can be estimated and adjusted for, while moment-adjusted imputation allows for the measurement error variance to be subject-specific, but it must be known or well-estimated.

In moment reconstruction, the aim is to create an observable random variable, $\mathbf{X}^* = \mathbf{X}^*(\mathbf{Q}, \mathbf{Z}, Y)$ with the same first two moments as those of $\mathbf{X}$ given (Y, $\mathbf{Z}$). In moment-adjusted imputation, the construction also requires that $\mathbf{U} = \text{Normal}(0, \boldsymbol{\Sigma}_u)$, and it aims to create a variable $\mathbf{X}^* = \mathbf{X}^*(\mathbf{Q}, \mathbf{Z}, Y)$, *a function of the observed data* $\mathbf{Q}$, $\mathbf{Z}$ and Y, that has multiple moments that are the same as those of $\mathbf{X}$, and also the same covariance structure with (Y, $\mathbf{Z}$) as that of $\mathbf{X}$.

A major appeal of moment reconstruction and moment-adjusted imputation is the promise that once the derived variable $\mathbf{X}^*$ is created, it can be used in all subsequent analyses, without the need for redoing a measurement error analysis from scratch each time a different risk model is proposed. Indeed, Freedman et al. (2004) use the moment reconstructed variable in logistic regression, linear discriminant analysis and in constructing a classification tree, simultaneously. Additionally, for example, if $\mathbf{X}$ is scalar and Y is binary, one might wish to model the effect of $\mathbf{X}$ on Y in a logistic regression with $\mathbf{X}$ modeled as linear, via a simple B-spline, or, following the typical epidemiological convention, as a step function, with either pre-defined or estimated categories. Both moment reconstruction and moment-adjusted imputation are of course only approximate methods, but they have been shown to have good performance in a variety of areas. See, for example, (Freedman et al., 2004, 2008; Thomas et al., 2011, 2013).

The literature in the measurement error field is vast, with four books (Fuller, 1987; Carroll et al., 2006; Gustafson, 2004; Buonaccorsi, 2010). There is also a long history of nutrition having an impact on the measurement error field (Rosner et al., 1990; Spiegelman et al., 2000; Sugar et al., 2007; Nusser et al., 1996; Carriquiry, 2003). There is a smaller set of papers on measurement error and radiation exposure (Prentice, 1982; Reeves et al., 1998; Kopecky et al., 2011; Pierce et al., 2009), which has classical measurement errors and Berkson errors (Reeves et al., 1998; Mallick et al., 2002; Delaigle et al., 2006; Schennach, 2013).

Our purpose is to extend moment reconstruction and moment-adjusted imputation from the classical, homoscedastic measurement error model. The problem we study is the case wherein the crucial predictor $\mathbf{X}_i$ arises as follows. For estimable parameters $\boldsymbol{\Theta}$, and individual-level random effects $\boldsymbol{\zeta}$ the distribution of which depends on parameters in $\boldsymbol{\Theta}$, for a known function $\mathcal{G}(\cdot)$, and for a sample with $i = 1, \ldots, n$, the crucial but unobserved true predictor satisfies

$$\mathbf{X}_i = \mathcal{G}(\mathbf{Q}_i, \boldsymbol{Z}_i, \boldsymbol{\Theta}, \boldsymbol{\zeta}_i). \quad (1)$$

The purpose of this note is to develop a method for model (1) that allows use of moment reconstruction and moment-adjusted imputation.

As stated, model (1) has no context. We justify it in detail in the context of dietary patterns research, an important field in nutritional surveillance and epidemiology, see Section 2. However, it is of much broader applicability to problems in biology, physics, or social science in which individual true exposures are modeled as a possibly complex function of mismeasured variables, covariates and random effects. For example, as detailed in Supplementary Material Appendix S.1, in radiation research (Reeves et al., 1998; Ostrouchov et al., 2000; Davis et al., 2002; Mallick et al., 2002), it is thought that calculated doses contain a mixture of Berkson errors (from physical transport systems modeling) and classical error (from measuring factors such a milk). If one knows the mixture percentage, the model in Reeves et al. (1998) and Mallick et al. (2002) is a special case of (1). If the mixture percentage is unknown and varies at the individual level, that too is a special case of (1).

We are motivated by the study of colorectal cancer Y in the National Institutes of Health-AARP Diet and Health Study (NIH-AARP) (Schatzkin et al., 2001; Reedy et al., 2008), with one of the risk predictors being the Healthy Eating Index-2005 (HEI-2005), (Guenther et al., 2008), a multi-component index meant to measure adherence to the 2005 USDA Dietary Guidelines for Americans. As described in Section 2, the HEI-2005 has a complex, error structure of the type embodied by (1). Our aim is to derive methods in the same vein as moment reconstruction and moment-adjusted imputation, but in this very different context.

In Section 2, we give a brief review of the HEI-2005 and the NIH-AARP study, and show how it fits into the form (1). Section 3 gives our basic approach, which in effect redefines the problem as a classical error problem, and describes implementation of moment

reconstruction and moment-adjusted imputation. Section 4 gives a data analysis of the NIH-AARP study, Section 5 gives simulation results, and Section 6 has concluding remarks. Supplementary Material contains additional results and details.

2. The HEI-2005 and the NIH-AARP Study

2.1 The NIH-AARP Study

In this section, we describe how model (1) can arise in dietary patterns research in the important fields of nutritional surveillance and epidemiology.

As described in Section 1, our main example is taken from the National Institutes of Health-AARP Diet and Health Study (NIH-AARP), with the outcome Y being an indicator of incident colorectal cancer for logistic regression; for Cox regression, Y is time until colorectal cancer. We did separate analyses for men and women, in the latter case deleting those with missing menopausal hormone therapy status, none of whom developed colorectal cancer. In the main study, the sample sizes were $n = 293{,}615$ for men and $n = 198{,}245$ for women. There were 2,151 men and 959 women who developed colorectal cancer. The covariates $\mathbf{Z}$ used were the same as in Reedy et al. (2008), consisting of age and dummy variable categories for education, ethnicity, body mass index, smoking status and physical activity. A food frequency questionnaire (FFQ) $\mathbf{Q}$ was obtained from all study participants.

The FFQ is known to be biased for usual nutritional intakes and also heteroscedastic, so that moment reconstruction and moment-adjusted imputation are not applicable for it. However, the NIH-AARP study has a small sub-study, known as a calibration study, in which 866 men and 854 women completed two 24-hour recalls (24HR). These recalls are assumed to be unbiased for usual intake, although heteroscedastic. We will use this sub-study to model usual intakes, resulting in a model that is a special case of (1).

2.2 HEI-2005

The Healthy Eating Index-2005 includes ratios of interrelated dietary components to energy and comprises 12 distinct component scores and a total summary score. Intakes of each food or nutrient, represented by one of the 12 components, are expressed as a ratio to energy intake, assessed, and ascribed a score. See Table 1 for a list of these components and the standards for scoring, and (Guenther et al., 2008) for details. The 12 HEI-2005 components represent 6 episodically consumed food groups (total fruit, whole fruit, total vegetables, dark green and orange vegetables and legumes or DOL, whole grains and milk), which are not consumed daily by most, 3 daily-consumed food groups (total grains; meat and beans; and oils), and 3 other daily-consumed dietary components (saturated fat; sodium; and calories from solid fats, alcoholic beverages and added sugars, or SoFAAS). The crucial statistical aspect of the data is that 24HR-reports for six episodically consumed food groups are zero-inflated, namely total fruit, whole fruit, whole grains, total vegetables, DOL and milk.

The short-term dietary instruments used, the two24-hour recalls, are assumed to be unbiased measures of usual dietary intake on the original scale. However, they are not homoscedastic, so that the classical measurement error model does not hold for them. In any case, as described in Section 2.1, they are not available for the main NIH-AARP study. In addition,

as seen in Table 1, the HEI-2005 total score is a highly nonlinear function of usual intakes. Zhang et al. (2011), see also Kipnis et al. (2009), use the assumption of unbiasedness of 24HR reports to model true usual intakes of episodically consumed dietary components. Zhang et al. (2011) show how the multivariate extension of this model could be specified for the true HEI-2005 and fit using MCMC.

It is not at all obvious that this problem is a special case of (1). Both Kipnis et al. (2009) and Zhang et al. (2011) show that it is, but only implicitly and not in the form of (1). In Appendix 6, we show this explicitly, although the modeling details of how the sub-study with the two 24-hour recalls is used is quite technical, and full details of this modeling process are described in Zhang et al. (2011).

3. Methods

3.1 Basic Approach

Our basic approach to constructing analogues of moment reconstruction and moment-adjusted imputation is to transform the data into a form amenable for these methods. We use of the Box-Cox transformation here, but the method can easily be adapted to include other transformations to normality, see Nusser et al. (1996) as one example. Define **W** as the vector of FFQ measurements for the total score and energy: it is of course a function of **Q**. We require that **X** and **W** have the same number of components p_x; in our case, $p_x = 2$. For a vector of parameters $\boldsymbol{\lambda}$ of length p_x, let $g(\mathbf{X}, \boldsymbol{\lambda})$ be the componentwise Box-Cox transformations for 1− Total Score/100 and Energy/2500, the former because the total score is left skew. We first assume that there are parameters $(\boldsymbol{\lambda}_w, \boldsymbol{\lambda}_x)$ such that

$$g(\mathbf{W}, \boldsymbol{\lambda}_w) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1^{\mathrm{T}} g(\mathbf{X}, \boldsymbol{\lambda}_x) + \boldsymbol{\beta}_2^{\mathrm{T}} \boldsymbol{Z} + \mathbf{U}, \quad (2)$$

where $\boldsymbol{\beta}_1$ is of full rank and **U** = Normal (0, $\boldsymbol{\Sigma}_u$) is independent of (Y, **X, Z**).

Define $\tilde{\mathbf{U}} = (\boldsymbol{\beta}_1^{\mathrm{T}})^{-1}\mathbf{U}$, is independent of **Z** and has covariance matrix $\boldsymbol{\Sigma}_{\tilde{u}}$. Define $\tilde{\mathbf{X}}$ = g(**X,** $\boldsymbol{\lambda}_x$), so that

$$\mathbf{W}_{\mathrm{mr}} = (\boldsymbol{\beta}_1^{\mathrm{T}})^{-1}\{g(\mathbf{W}, \boldsymbol{\lambda}_w) - \boldsymbol{\beta}_0 - \boldsymbol{\beta}_2^{\mathrm{T}} \boldsymbol{Z}\} = \tilde{\mathbf{X}} + \tilde{\mathbf{U}}, \quad (3)$$

Finally, with **V** = Normal(0, $\boldsymbol{\Sigma}_x$) is independent of (**Z, U**), we assume that

$$g(\mathbf{X}, \boldsymbol{\lambda}_x) = \boldsymbol{\alpha}_0^{\mathrm{T}} + \boldsymbol{\alpha}_1^{\mathrm{T}} \boldsymbol{Z} + \mathbf{V}. \quad (4)$$

With this construction, we now have a scenario where moment reconstruction and moment-adjusted imputation can be applied directly, since (3) is a classical measurement error model. Of course, the classical model (3) is defined on the transformed scale $g(X, \lambda_x)$, so when using it as a linear predictor, the implicit assumption is that the risk model would be fit on

this transformed scale, which may or not be fit the data. This is an issue for any transformation of predictors.

3.2 Estimating the Parameters in Section 3.1

Remember that **W** is a function of **Q**. We use Appendix equations (A.5)–(A.6) and equations (2)–(3). Estimation of the transformation parameters ($\boldsymbol{\lambda}_w$, $\boldsymbol{\lambda}_x$) is required. In general, this would not be possible since we do not observe (**Z, Q, W, X**) even on a subset of the data. However, since $\boldsymbol{\zeta}$ is independent of (**Z, Q, W**), by generating realizations of $\boldsymbol{\zeta}$ and substituting into (A.6), we can generate (**Z, Q, W, X***) that have the same joint distribution as (**Z, Q, W, X**). Parameter estimates can therefore be obtained easily from these simulated random variables. In addition, this allows us to check models (2) and (4).

We outline here how the transformation parameters were estimated using said pairs. Let $\hat{\boldsymbol{\alpha}}_0(\boldsymbol{\lambda}_x)$ and $\hat{\boldsymbol{\alpha}}_1(\boldsymbol{\lambda}_x)$ denote the least squares parameter estimates when performing a linear regression of covariates $g(\mathbf{X}^*, \boldsymbol{\lambda}_x)$ on **Z** for a fixed value of $\boldsymbol{\lambda}_x$. Define residuals $\mathbf{V}^*(\boldsymbol{\lambda}_x) = g(\mathbf{X}^*, \boldsymbol{\lambda}_x) - \hat{\boldsymbol{\alpha}}_0(\boldsymbol{\lambda}_x) - \hat{\boldsymbol{\alpha}}_1(\boldsymbol{\lambda}_x)\mathbf{Z}$. Since the distribution of $\mathbf{V}^*$ is assumed Gaussian for the true value of $\boldsymbol{\lambda}_x$, the estimated transformation parameter $\hat{\boldsymbol{\lambda}}_x$ is, component-wise, the value that maximizes the absolute correlation between the percentiles of $\mathbf{V}^*$ and the percentiles of the standard Gaussian distribution. A similar procedure is used to estimate $\hat{\boldsymbol{\lambda}}_w$, and then the other parameters.

3.3 Moment-Adjusted Imputation

Model (3) is exactly a classical measurement error model, to which moment-adjusted imputation can be applied. In principle, one has to do a bivariate moment-adjusted imputation, for which programs are not yet available. However, in our context, the HEI-2005 total score is very nearly independent of energy intake, and thus for simplicity we used the programs mentioned in Thomas et al. (2011) separately for each variable. A multivariate extension to MAI has recently been considered by Thomas et al. (2013). The latter would be preferable when strong dependence exists between the variables under consideration.

3.4 Moment Reconstruction

Let $\{\mathrm{cov}(\mathbf{W}_{\mathrm{mr}}|Y, \mathbf{Z})\}^{1/2}$ be the symmetric square root of $\mathrm{cov}(\mathbf{W}_{\mathrm{mr}}|Y, \mathbf{Z})$. Define $m(Y, \mathbf{Z}) = E(\mathbf{W}_{\mathrm{mr}}|Y, \mathbf{Z})$ and $G(Y, \mathbf{Z}) = \{\mathrm{cov}(\mathbf{W}_{\mathrm{mr}}|Y, \mathbf{Z})^{1/2}\}^{-1}\{\mathrm{cov}(\tilde{\mathbf{X}}|Y, \mathbf{Z})\}^{1/2}$. Moment reconstruction now proceeds by substituting the unobserved $\tilde{\mathbf{X}}$ by $\mathbf{X}_{\mathrm{mr}} = m(Y, \mathbf{Z})\{I - G(Y, \mathbf{Z})\} + \mathbf{W}_{\mathrm{mr}}G(Y, \mathbf{Z})$ which has been constructed so that the first two conditional moments of $\tilde{\mathbf{X}}$ and $\mathbf{X}_{\mathrm{mr}}$ are equal. Of course, to get to this point, the additional parameters ($\boldsymbol{\lambda}_w$, $\boldsymbol{\lambda}_x$) and ($\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\alpha}_0$, $\boldsymbol{\alpha}_1$) also need to be estimated, see Section 3.1. The model of interest is assumed to be a function of $\tilde{\mathbf{X}}$, with the function known up to a vector of parameters.

In any given example of moment reconstruction, constructing $m(Y, \mathbf{Z})$ and $G(Y, \mathbf{Z})$ is done on a case-by-case basis. Freedman et al. (2004) show how to do this explicitly if there are no additional covariates **Z**, and if Y is binary as in logistic regression. Specifically, $m(Y, \mathbf{Z})$ is the mean of $\tilde{\mathbf{X}}$ among those sharing the same values of Y, and $\mathrm{cov}(\tilde{\mathbf{X}}|Y, \mathbf{Z})$ is the covariance of $\mathbf{W}_{\mathrm{mr}}$ among those sharing the same values of Y minus $\mathrm{cov}(\tilde{\mathbf{U}})$. In the example of Section 4, however, **Z** is of dimension > 20 and thus this simple device is not applicable. Instead, one

can perform linear regressions of $\hat{\mathbf{X}}^*_{\rm mr}$ on **Z** separately for the cases (Y = 1) and controls (Y = 0), from which $m(Y, \mathbf{Z})$ is estimated directly, as is cov($\mathbf{W}_{\rm mr}|\mathbf{Z}$), and then cov($\mathbf{X}_{\rm mr}|Y, \mathbf{Z}$) = cov($\mathbf{W}_{\rm mr}|Y, \mathbf{Z}$) − cov($\tilde{\mathbf{U}}$).

In our problem, with many covariates, we used the following device for logistic regression. Using the parameters estimates found in Section 3.2, define $\hat{\mathbf{X}}^*_{\rm mr}=g(\mathbf{X}^*,\hat{\boldsymbol{\lambda}}_x)$ and $\hat{\mathbf{W}}_{\rm mr}=(\hat{\boldsymbol{\beta}}_1^{\rm T})^{-1}\{g(\mathbf{W},\hat{\boldsymbol{\lambda}}_w)-\hat{\boldsymbol{\beta}}_0-\hat{\boldsymbol{\beta}}_2^{\rm T}Z\}$. The estimate $\hat{m}(Y, \mathbf{Z})$ of $m(Y, \mathbf{Z}) = E(\mathbf{W}_{\rm mr}|Y, \mathbf{Z})$ is found by performing separate linear regressions of $\hat{\mathbf{X}}^*_{\rm mr}$ on the covariates **Z** for both the cases (Y = 1) and controls (Y = 0). In estimating the covariance component, we assume that cov ($\mathbf{W}_{\rm mr}|Y, \mathbf{Z}$) = cov($\tilde{\mathbf{X}}|Y, \mathbf{Z}$) + cov($\tilde{\mathbf{U}}$). We also assume that cov ($\mathbf{W}_{\rm mr}|Y, \mathbf{Z}$) only depends on **Z** through). $m(Y, \mathbf{Z})$. While we are unable to estimate cov ($\mathbf{X}_{\rm mr}|Y, \mathbf{Z}$) directly from the data, we are able to find estimates of both cov ($\mathbf{W}_{\rm mr}|Y, \mathbf{Z}$) and cov($\tilde{\mathbf{U}}$). Define residuals $\tilde{\mathbf{U}}_{i,\rm res}=(\hat{\boldsymbol{\beta}}_1^{\rm T})^{-1}\{g(\mathbf{W}_i,\hat{\boldsymbol{\lambda}}_w)-\hat{\boldsymbol{\beta}}_0-\hat{\boldsymbol{\beta}}_2^{\rm T}Z_i\}-g(\mathbf{X}_i,\hat{\boldsymbol{\lambda}}_x)$ and let $\hat{\boldsymbol{\Sigma}}_{\tilde{u}}$ be the sample covariance matrix of the $\tilde{\mathbf{U}}_{i,\rm res}$, the estimate of cov($\tilde{\mathbf{U}}$). The estimate $\widehat{\rm cov}\ (\mathbf{W}_{\rm mr}|Y,Z)$ is found by calculating $\hat{\mathbf{X}}^*_{\rm mr}-\hat{m}(Y,Z)$ in both the cases and controls, and then finding the covariance matrices corresponding to those residuals.

3.5 Modified Regression Calibration

Regression calibration is defined as replacing a latent variable by its expectation given the observed covariates. We do this in the original data scale, as follows. We use the characterization $\mathbf{X} = \mathscr{G}(\mathbf{Q}, \mathbf{Z}, \boldsymbol{\Theta}, \boldsymbol{\Sigma}_\zeta, \boldsymbol{\Sigma}_\xi, \boldsymbol{\zeta})$ given in (A.6). We compute $E(\mathbf{X}|\mathbf{Q}, \mathbf{Z})$ by Monte-Carlo. Set B = 500, and generate ($\boldsymbol{\zeta}_{1,\rm RC}$, …, $\boldsymbol{\zeta}_{B,\rm RC}$) = Normal(0, $\boldsymbol{\Sigma}_\zeta$). Let the estimates of ($\boldsymbol{\Theta}, \boldsymbol{\Sigma}_\zeta, \boldsymbol{\Sigma}_\xi$) be ($\hat{\boldsymbol{\Theta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \hat{\boldsymbol{\Sigma}}_\xi$). Then $\hat{E}(\mathbf{X}|\mathbf{Q},Z)=\hat{\mathbf{X}}_{\rm RC}=B^{-1}\sum_{b=1}^{B}\mathscr{G}(\mathbf{Q},Z,\hat{\boldsymbol{\Theta}},\hat{\Sigma}_\varsigma,\hat{\Sigma}_\xi,\varsigma_{b,\rm RC})$

This procedure is completely standard when $\mathbf{X} = (X_T, X_E)$ enters the risk model linearly. However, when we do regression of fixed intervals, or use a B-spline basis for usual total score, a modification is required. In both those cases, **X** enters the model linearly through functions $\{g_1(\mathbf{X}), g_2(\mathbf{X}), \ldots, g_K(\mathbf{X})\}$. For example, in the quintile analysis, K = 6 and $g_1(\cdot)$, …,$g_5(\cdot)$ are the indicators that X_T are in the first through fifth quintile, while $g_6(\mathbf{X}) = X_E$, and similarly for the B-spline analysis. In such cases, we replace $g_k(\mathbf{X})$ by $E\{g_k(\mathbf{X})|\mathbf{Q}, \mathbf{Z}\}$: for the quintile analysis, see equation (**??**) in Appendix S.2 of the Supplementary Material.

4. The NIH-AARP Study Analysis

4.1 Overview

The data are described in Section 2.1. We fit the data using logistic regression: results were very similar for Cox regression, where, following Thomas et al. (2011), **Z** was augmented by case-control status.

- We did 5 different analyses with 3 different models. The analyses were (a) use of the FFQ in the original scale and ignoring measurement error; (b) regression calibration (RC) on the original scale as described in Section 3.5; (c) moment reconstruction (MR); (d) moment-adjusted imputation

(MAI); and (e) Monte-Carlo maximum likelihood (MCML) on the original scale, with the score functions computed using $B = 500$ simulations, see Appendix S.2 of the Supplementary Material for details.

- The 3 different models were (i) linear logistic regression; (ii) because there was some hint of curvature in the regression model when using the FFQ, we also fit a quadratic B-spline with 4 basis functions; and (iii) dummy variable regression for the HEI-2005 total score based on the estimated quintiles of the true total score. For (iii), because all transformations are monotone, the quintiles in the transformed scale are immediate. For men, the quintile break points are (50.6, 58.0, 64.0, 70.3), while for women they are (55.9, 62.9, 68.3, 73.7).
- When evaluating (i) and (ii), we computed the relative risk when moving from a true total score of 45, representing a rather poor diet, to a true total score of 75, representing a good diet. For women, these represent the 5^{th} and 92^{nd} percentiles of usual intake total score, while they are the 10^{th} and 91^{st} percentiles for men. In addition, these two values cover most of the range for the FFQ and regression calibration expectations.

When evaluating (iii), we computed relative risk between the first and fifth quintile, also representing a change from a poor diet to a good diet.

The Box-Cox transformation parameters for moment reconstruction and moment adjusted imputation were 0.66 for men and 0.40 for women. For energy, they were 0.23 and 0.36, respectively. Standard errors were estimated using 500 bootstrap data sets.

4.2 Results

Results for the analysis of the HEI-2005 Total Score are provided in Table 2. In the Supplementary Material Figures 1–2, we display the Bspline fits by the various methods.

Consider first the analysis for men. With one exception, discussed below, the relative risks are consistent within method. For the linear risk model and the spline model, moment-adjusted imputation, regression calibration and Monte-Carlo maximum likelihood all have risks about 10% lower than those estimated by the FFQ, with moment reconstruction between the first three methods and the FFQ. The only anomaly arises in the quintile analysis, where Monte-Carlo maximum likelihood estimates a relative risk 16% smaller than that of the FFQ, and about 15% smaller than for moment-adjusted imputation. The quintile model actually does not fit the data at all well, and this may reflect that had $\mathbf{X}$ been observable, a quintile analysis would have suggest much more attenuation of risk when using the FFQ compared to the linear model.

The results for women are interesting. We do not observe the same phenomenon about the quintile analysis using Monte-Carlo maximum likelihood as was observed in men. The spline model does appear more appropriate than a linear model (Figures 1–2 in the Supplementary Material), and if we look at the spline model results, all the measurement error corrections suggest a large attenuation of risk when using the FFQ. Perhaps of most

interest is that there is no statistically significant effect of HEI-2005 on colorectal cancer when using the FFQ. However, all the measurement error correction methods are different, with p-values ranging from 0.0% to 5.1%. This may seem paradoxical, since the folklore is that measurement error can be ignored when testing null effects, but as discussed in Chapter 10 of Carroll et al. (2006), such folklore is generally true only if there are no covariates measured without error that are also correlated with **X**. In our case, there are over 20 covariates **Z** in the risk model, and, importantly, those covariates are also predictors of **X** in the model of Zhang et al. (2011) discussed in Section 6, and in fact diet composition does depend on the demographic factors making up **Z**. We believe it is this phenomenon that leads to the change from non-statistical significance to statistical significance in the women.

5. Simulation Study

To simulate data that has properties similar to the observed data, several steps are necessary. First, one needs to simulate a calibration data set (usual intake). The calibration data requires specification of model parameters (Θ, $\mathbf{\Sigma}_\zeta$, $\mathbf{\Sigma}_\xi$), which are estimated in Zhang et al. (2011). For the purpose of this simulation, we used these aforementioned estimated values as the true model parameters. Given simulated usual intake, one can simulate total score and energy, which are necessary to calculate the risk function associated with colorectal cancer and therefore simulate this outcome. In this simulation study, two different risk functions are considered. Let $H(x) = \{1 + \exp(-x)\}^{-1}$ be the logistic distribution function, and

$$\text{pr}\ (Y=1|\mathbf{X}, \mathbf{Q}, \boldsymbol{Z}, \mathbf{U})=H(\gamma_0+\mathbf{X}^T\gamma_1+\boldsymbol{Z}^T\gamma_2).$$

The risk functions considered are respectively linear (**X** includes total score linearly) and a quintile function (**X** includes a step function based on the quintiles of total score). It is then possible to apply the different methods discussed here (MR, MAI, RC, MCML) to use the intake observations with measurement error present to estimate the relative risk associated with an increase in total score. When the specified risk function is linear, both a linear and quintile model are fit, while when the specified risk function is quintile, only a quintile model is fit. In each instance, 500 data sets were generated and the relative risk (RR) was estimated. Table 3 provides a summary of the average RR from the 500 simulations and the standard deviation of the estimated RR.

In summary, none of the methods show serious bias, and moment-adjusted imputation and moment reconstruction are comparable in performance, although moment-adjusted imputation estimates tend to have smaller standard deviation among females.

6. Discussion

Moment reconstruction and moment-adjusted imputations, as well as regression calibration, are methods that, up to now, have only been applied in the classic measurement error setting. We have shown how to apply these methods for non-classical measurement error structures in our context in which exposure is modeled by a complex, nonlinear random effects modeling process in physics, biology or social science applications. While these methods are

only approximations, we have shown that they perform well in the context of the HEI-2005 problem.

In the present setting, our method relied on the availability of a calibration sample, from which the covariance matrices of the random effects and measurement error components could be estimated. This is typical in measurement error problems, where one has to assume that certain variance components are known when replicate measurements are not available.

For the colorectal cancer application discussed in this paper, moment reconstruction and moment-adjusted imputation performed well for both a quantile model, which is often used in epidemiological models, as well as a spline model which allows for some flexibility. The results compared favorably to a Monte Carlo maximum likelihood approach. The latter approximates the exact solution, but is computationally intensive.

The simulations and data analysis show give the impression that regression calibration is similar to moment reconstruction and moment assisted imputation. This is not always the case. In Section S.4 of the Supplementary Material, we give results of a simulation where it is shown that two forms of regression calibration which uses dummy variables can have substantial bias. Freedman et al. (2004) and Freedman et al. (2008) compare moment reconstruction and ordinary regression calibration and efficient regression calibration (Spiegelman et al., 2001) in the classical additive, homoscedastic measurement error model. They find that moment reconstruction dominates ordinary regression calibration, while if there is a calibration sub-study that also has the response, then enhanced regression calibration dominates moment reconstruction when the correlation between Y and X is modest, but not necessarily otherwise. Thomas et al. (2011) and Thomas et al. (2013) do simulations with nondifferential measurement error and conclude that moment-adjusted imputation dominates moment reconstruction. However, moment reconstruction is easily applied if the measurement error is differential, as it might be in a case-control study.

Finally, model (2) has an interesting feature in that it can also be thought of as a complex, non-standard, nonlinear and heteroscedastic Berkson model. Define $\mathbf{W}_{i*} = E\{\mathscr{G}(\mathbf{Q}_i, \mathbf{Z}_i, \boldsymbol{\Theta}, \boldsymbol{\zeta}_i)|\mathbf{Q}_i, \mathbf{Z}_i\}$. Then $\mathbf{X}_i = \mathbf{W}_{i*} + \upsilon_i$, where $\upsilon_i = \mathscr{G}(\mathbf{Q}_i, \mathbf{Z}_i, \boldsymbol{\Theta}, \boldsymbol{\zeta}_i) - \mathbf{W}_{i*}$. This looks like a standard Berkson model, since $E(\upsilon_i|\mathbf{Q}_i, \mathbf{Z}_i) = 0$, but υ_i is heteroscedastic. However, our methods do not rely on thinking about this as a Berkson model.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research of Potgieter, Wei and Carroll was supported by a grant from the National Cancer Institute (U01-CA057030).

References

Buonaccorsi, JP. Measurement Error: Models, Methods and Applications. Chapman & Hall; 2010.

Carriquiry AL. Estimation of usual intake distributions of nutrients and foods. The Journal of Nutrition. 2003; 133:601S–608S. [PubMed: 12566510]

Carroll, RJ.; Ruppert, D.; Stefanski, LA.; Crainiceanu, CM. Measurement Error in Nonlinear Models: A Modern Perspective, Second Edition. Chapman and Hall; 2006.

Davis, S.; Kopecky, KJ.; Hamilton, TE.; Onstad, L.; King, B.; Saporito, M.; Callahan, C. Hanford Thyroid Disease Study: Final Report. 2002. http://www.cdc.gov/nceh/radiation/hanford/htdsweb.

Delaigle A, Hall P, Qiu P. Nonparametric methods for solving the Berkson errors-in-variables problem. Journal of the Royal Statistical Society: Series B. 2006; 68:201–220.

Freedman LS, Feinberg V, Kipnis V, Midthune D, Carroll RJ. A new method for dealing with measurement error in explanatory variables of regression models. Biometrics. 2004; 60:171–181.

Freedman LS, Midthune D, Carroll RJ, Kipnis V. A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. Statistics in Medicine. 2008; 27:5195–6216. [PubMed: 18680172]

Fuller, WA. Measurement Error Models. Chapman and Hall/CRC; 1987.

Guenther PM, Reedy J, Krebs-Smith SM, Reeve BB. Evaluation of the Healthy Eating Index-2005. Journal of the American Dietetic Association. 2008; 108:1854–1864. [PubMed: 18954575]

Gustafson, P. Measurement Error and Misclassication in Statistics and Epidemiology. Chapman and Hall/CRC; 2004.

Kipnis V, Midthune D, Buckman DW, Dodd KW, et al. Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics. 2009; 65:1003–1010. [PubMed: 19302405]

Kopecky KJ, Stepaneko V, Rivikind N, et al. Childhood thyroid cancer, radiation dose from Chernobyl and dose uncertainties in Bryansk Oblast, Russia: A population-based case-control study. Radiation Research. 2011; 166:367–374.

Mallick B, Hoffman FO, Carroll RJ. Semiparametric regression modeling with mixtures of berkson and classical error, with application to fallout from the nevada test site. Biometrics. 2002; 58:13–20. [PubMed: 11890308]

Nusser SM, Carriquiry AL, Dodd KW, Fuller WA. A semiparametric transformation approach to estimating usual daily intake distributions. Journal of the American Statistical Association. 1996; 91:1440–1449.

Ostrouchov G, Frome EL, Kerr GD. Dose estimation from daily and weekly dosimetry data. ORNL/TM-1999/282, Oak Ridge National Laboratory, Oak Ridge, Tennessee. 2000

Pierce DA, Væth M, Cologne JB. Allowance for random dose estimation errors in atomic bomb survivor studies: A revision. Radiation Research. 2009; 170:118–126.

Prentice RL. Covariate measurement errors and parameter estimation in a failure time regression model. Biometrika. 1982; 69:331–342.

Reedy JR, Mitrou PN, Krebs-Smith SM, Wirfält E, Flood AV, Kipnis V, et al. Index-based dietary patterns and risk of colorectal cancer: the nih-aarp diet and health study. American Journal of Epidemiology. 2008; 168:38–48. [PubMed: 18525082]

Reeves G, Cox D, Darby S, Whitley E. Some aspects of measurement error in explanatory variables for continuous and binary regression models. Statistics in Medicine. 1998; 17:2157–2177. [PubMed: 9802176]

Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. American Journal of Epidemiology. 1990; 132:734–745. [PubMed: 2403114]

Schatzkin A, Subar AF, Thompson FE, et al. Design and serendipity in establishing a large cohort with wide dietary intake distributions: the national institutes of health-aarp diet and health study. American Journal of Epidemiology. 2001; 154:1119–1125. [PubMed: 11744517]

Schennach SM. Regressions with Berkson errors in covariatesa nonparametric approach. Annals of Statistics. 2013; 41:1642–1668.

Spiegelman D, Carroll RJ, Kipnis V. Efficient regression calibration for logistic regression in main study/internal validation study designs with an imperfect reference instrument. Statistics in Medicine. 2001; 20:139–160. [PubMed: 11135353]

Spiegelman D, Rosner B, Logan R. Estimation and inference for logistic regression with covariate misclassification and measurement error in main study/validation study designs. Journal of the American Statistical Association. 2000; 95:51–61.

Sugar EA, Wang C-Y, Prentice RL. Logistic regression with exposure biomarkers and flexible measurement error. Biometrics. 2007; 63:143–151. [PubMed: 17447939]

Thomas L, Stefanski LA, Davidian M. A moment-adjusted imputation method for measurement error models. Biometrics. 2011; 67:1461–1470. [PubMed: 21385161]

Thomas L, Stefanski LA, Davidian M. Moment adjusted imputation for multivariate measurement error data with applications to logistic regression. Computational Statistics & Data Analysis. 2013; 67:15–24. [PubMed: 24072947]

Zhang S, Midthune D, Guenther PM, Krebs-Smith SM, Kipnis V, Dodd KW, Buckman DW, Tooze JA, Freedman LS, Carroll RJ. A new multivariate measurement error model with zero-inflated dietary data, and its application to dietary assessment. Annals of Applied Statistics. 2011; 5:1456–1487. [PubMed: 21804910]

Zhang S, Midthune D, P'erez A, Buckman DW, Kipnis V, Freedman LS, Dodd KW, Krebs-Smith SM, Carroll RJ. Fitting a bivariate measurement error model for episodically consumed dietary components. International Journal of Biostatistics. 2011; 7 Available at: http://www.bepress.com/ijb/vol7/iss1/1.

Appendix: Sketch of Technical Arguments

The Model in Zhang et al. (2011)

The purpose of this section is to show that model (1) is justified for the NIH-AARP Study. Because of space constraints we provide only a sketch of that method.

Using the repeated 24HR recall data, for each of the episodically consumed food groups, two variables are defined: (a) whether a food from that group was consumed; and (b) the amount of the food that was reported on the 24HR recall. For the 6 daily-consumed food groups and nutrients, only one variable indicating the consumption amount is defined. In addition, the amount of energy that is calculated from the 24HR recall is of interest. The number of dietary variables for each 24HR recall is thus 12+6+1=19, each collected at 2 time points. The observed data are R_{ijk} for the i^{th} person, the j^{th} variable and the k^{th} replicate, $j = 1, \ldots, 19$ and $k = 1, 2$. Set $\mathbf{R}_{ik} = (R_{i1k}, \ldots, R_{i,19,k})^{\mathrm{T}}$, where (a) $R_{i,2\ell-1,k}$ = Indicator of whether dietary component # ℓ is consumed, with $\ell = 1, 2, 3, 4, 5, 6$; (b) $R_{i,2\ell,k}$ = Amount of food # ℓ consumed. This equals zero, of course, if none of food #ℓ is consumed, with $\ell = 1, 2, 3, 4, 5, 6$; (c) $R_{i,\ell+6,k}$ = Amount of non-episodically consumed food or nutrient #ℓ, with $\ell = 7, 8, 9, 10, 11, 12$; and (d) $R_{i,19,k}$ = Amount of energy consumed as reported by the 24HR recall.

Each of the 6 episodically consumed foods has 2 sets of latent variables, one for consumption and one for amount, while the 6 daily-consumed foods and nutrients as well as energy have 1 latent variable each, for a total of 19. The latent random variables are $\boldsymbol{\xi}_{ik} = (\xi_{i1k}, \ldots, \xi_{i,19,k}) = \text{Normal}(0, \boldsymbol{\Sigma}_{\xi})$ and $\boldsymbol{\zeta}_i = (\zeta_{i1}, \ldots, \zeta_{i,19}) = \text{Normal}(0, \boldsymbol{\Sigma}_{\zeta})$, which are mutually independent. Technically, $\boldsymbol{\Sigma}$ is a patterned covariance matrix, see Zhang et al. (2011) for details. As before, $\mathbf{Z}$ represents covariates while $\mathbf{Q}$ represents the food frequency questionnaire. In this model, food $\ell = 1, \ldots, 6$ being consumed on day k is equivalent to observing the binary $R_{i,2\ell-1,k}$, where

$$R_{i,2\ell-1,k}=1 \iff S_{i,2\ell-1,k}=(1,\mathbf{Q}_i^{\mathrm{T}},\boldsymbol{Z}_i^{\mathrm{T}})\boldsymbol{\theta}_{2\ell-1}+\zeta_{i,2\ell-1}+\xi_{i,2\ell-1,k}>0. \quad \text{(A.1)}$$

Define the Box-Cox transformation as $g(y, \lambda) = (y^{\lambda} - 1)/\lambda$ for $\lambda \neq 0$ and $= \log(y)$ if $\lambda = 0$. If the food is consumed, we model the amount reported, $R_{i,2\ell,k}$, as

$$[g_{\mathrm{tr}}(R_{i,2\ell,k},\lambda_\ell)|R_{i,2\ell-1,k}=1]=S_{i,2\ell,k}=(1,\mathbf{Q}_i^{\mathrm{T}},\boldsymbol{Z}_i^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{\theta}_{2\ell}+\zeta_{i,2\ell}+\xi_{i,2\ell,k}, \quad \text{(A.2)}$$

where $g_{\mathrm{tr}}(y,\lambda)=\sqrt{2}\{g(y,\lambda)-\mu(\lambda)\}/\sigma(\lambda)$, $g(y, \lambda)$ is the usual Box-Cox transformation defined above with transformation parameter λ, and $\{\mu(\lambda), \sigma(\lambda)\}$ are the sample mean and standard deviation of $g(y, \lambda)$, computed from the nonzero food data. This standardization improves the numerical performance of the algorithm without affecting conclusions.

The reported consumption of daily consumed foods/nutrients plus energy is modeled as

$$g_{\mathrm{tr}}(R_{i,\ell+6,k},\lambda_\ell)=S_{i,\ell+6,k}=(1,\mathbf{Q}_i^{\mathrm{T}},\boldsymbol{Z}_i^{\mathrm{T}})\boldsymbol{\theta}_{\ell+6}+\zeta_{i,\ell+6}+\xi_{i,\ell+6,k}, \quad \text{(A.3)}$$

for $\ell = 7,\ldots, 13$. As seen in (A.2)–(A.3), different transformations $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{13})^{\mathrm{T}}$ are allowed to be used for the different types of dietary components.

Denote the collection of $\boldsymbol{\theta}_j$ as $\boldsymbol{\Theta}$. Zhang et al. (2011) use MCMC to estimate $(\boldsymbol{\Theta}, \boldsymbol{\Sigma}_\zeta, \Sigma_\xi)$. From that, usual intake and the usual HEI-2005 component score are defined as follows. Consider the first episodically consumed dietary component, a food group. Since the 24hour recalls are unbiased for a person's usual intake, the usual intake is the expectation of the reported intake conditional on the person's random effects $\boldsymbol{\zeta}_i$. Let $g_{\mathrm{tr}}^{-1}(\cdot)$ be the inverse transformation, and let $\Phi(\cdot)$ be the standard normal distribution function. Then, a person's usual intake of the first episodically consumed dietary component is

$$\begin{aligned}X_{i1,\mathrm{com}}&=X_{i1,\mathrm{com}}(\mathbf{Q}_i,\boldsymbol{Z}_i,\boldsymbol{\Theta},\boldsymbol{\Sigma}_\zeta,\boldsymbol{\Sigma}_\xi,\boldsymbol{\zeta}_i)=E(R_{i2}|\mathbf{Q}_i,\boldsymbol{Z}_i,\boldsymbol{\Theta},\boldsymbol{\Sigma}_\zeta,\boldsymbol{\Sigma}_\xi,\zeta_{i1},\zeta_{i2})\\&=\Phi\{(1,\mathbf{Q}_i^{\mathrm{T}},\boldsymbol{Z}_i^{\mathrm{T}})\boldsymbol{\theta}_1+\zeta_{i1}\}E[g_{\mathrm{tr}}^{-1}\{(1,\mathbf{Q}_i^{\mathrm{T}},\boldsymbol{Z}_i^{\mathrm{T}})\boldsymbol{\theta}_2+\zeta_{i2}+\xi_{i21},\lambda_i\}|\boldsymbol{\zeta}_i].\end{aligned} \quad \text{(A.4)}$$

Some fix-ups are used to make the expectation computable, but the details are not of interest here. Usual intake for the other episodically consumed food groups is defined similarly, and similarly for the daily consumed components, which do not have the leading term in (A.4).

Thus, for functions $\mathscr{G}_T$ and $\mathscr{G}_E$, the true HEI-2005 total score, X_T, and energy, X_E are

$$X_T=\mathscr{G}_T(\mathbf{Q},\boldsymbol{Z},\boldsymbol{\Theta},\boldsymbol{\Sigma}_\zeta,\boldsymbol{\Sigma}_\xi,\boldsymbol{\zeta}); \quad X_E=\mathscr{G}_E(\mathbf{Q},\boldsymbol{Z},\boldsymbol{\Theta},\boldsymbol{\Sigma}_\zeta,\boldsymbol{\Sigma}_\xi,\boldsymbol{\zeta}), \quad \text{(A.5)}$$

where $\boldsymbol{\zeta} = \text{Normal}(0, \boldsymbol{\Sigma}_\zeta)$ is independent of $(\mathbf{Z}, \mathbf{Q})$. Setting $(X_T, X_E)^{\mathrm{T}} = \mathbf{X}$, we write

$$\mathbf{X}=\mathscr{G}(\mathbf{Q},\boldsymbol{Z},\boldsymbol{\Theta},\boldsymbol{\Sigma}_{\zeta},\boldsymbol{\Sigma}_{\xi},\boldsymbol{\zeta})=\{\mathscr{G}_{T}(\mathbf{Q},\boldsymbol{Z},\boldsymbol{\Theta},\boldsymbol{\Sigma}_{\zeta},\boldsymbol{\Sigma}_{\xi},\boldsymbol{\zeta}),\mathscr{G}_{E}(\mathbf{Q},\boldsymbol{Z},\boldsymbol{\Theta},\boldsymbol{\Sigma}_{\zeta},\boldsymbol{\Sigma}_{\xi},\boldsymbol{\zeta})\}^{\mathrm{T}}, \quad \text{(A.6)}$$

which is the specific form of (1) for this application.

Table 1

Description of the HEI-2005 scoring system. Except for saturated fat and SoFAAS, density is obtained by multiplying usual intake by 1000 and dividing by usual intake of kilo-calories. For saturated fat, density is 9 × 100 usual saturated fat (grams) divided by usual calories, i.e., the percentage of usual calories coming from usual saturated fat intake. For SoFAAS, the density is the percentage of usual intake that comes from usual intake of calories, i.e., the division of usual intake of SoFAAS by usual intake of calories. Here, “DOL” is dark green and orange vegetables and legumes. Also, “SoFAAS” is calories from solid fats, alcoholic beverages and added sugars. The total HEI-2005 score is the sum of the individual component scores.

Component	Units	HEI-2005 score calculation
Total Fruit	cups	min (5, 5 × (density/.8))
Whole Fruit	cups	min (5, 5 × (density/.4))
Total Vegetables	cups	min (5, 5 × (density/1.1))
DOL	cups	min (5, 5 × (density/.4))
Total Grains	ounces	min (5, 5 × (density/3))
Whole Grains	ounces	min (5, 5 × (density/1.5))
Milk	cups	min (10, 10 × (density/1.3))
Meat and Beans	ounces	min (10, 10 × (density/2.5))
Oil	grams	min (10, 10 × (density/12))
Saturated Fat	% of	if density ≥ 15 score = 0
	energy	else if density ≤ 7 score = 10
		else if density > 10 score = 8 − (8 × (density − 10)/5)
		else, score = 10 − (2 × (density − 7)/3)
Sodium	milligrams	if density ≥ 2000 score=0
		else if density ≤ 700 score=10
		else if density ≥ 1100
		score = 8 − {8 × (density − 1100)/(2000 − 1100)}
		else score = 10 − {2 × (density − 700)/(1100 − 700)}
SoFAAS	% of	if density ≥ 50 score = 0
	energy	else if density ≤ 20 score=20
		else score = 20 − {20 × (density − 20)/(50 − 20)}

Table 2

Logistic regression analysis of the NIH-AARP Diet and Health Study for the HEI-2005 total score in Section 4. There are five methods considered: (a) moment reconstruction (MR); (b) moment-adjusted imputation (MAI), (c) regression calibration (RC); (d) the food frequency questionnaire (FFQ); and (e) Monte-Carlo maximum likelihood (MCML): the latter three were all done in the original data scale. Within each method the predictor either entered linearly (Lin), via quintiles (Quin) or via a B-spline (Spl). Displayed are the relative risk (Rel. Risk, in bold face), the p-value, and the lower and upper 95% confidence bounds (L 95% and H 95%, respectively) for the relative risk. The relative risk for the linear and spline analyses were the relative risk for moving from a total score of 45 to a total score of 75, while the relative risk for the quintile analysis was for the quintiles of the usual HEI-2005 total score. The quintile analysis for regression calibration is not included because it is known that categorization induces differential measurement error in regression calibration unless the true risk function is actually a step function of the categories.

		Men				Women			
		Rel. Risk	**p-value**	**L 95%**	**H 95%**	**Rel. Risk**	**p-value**	**L 95%**	**H 95%**
MR									
	Lin	**0.699**	< 0.001	0.614	0.796	**0.767**	0.005	0.637	0.925
	Quin	**0.710**	< 0.001	0.613	0.822	**0.790**	0.029	0.639	0.976
	Spl	**0.725**	< 0.001	0.634	0.830	**0.729**	0.002	0.600	0.887
MAI									
	Lin	**0.652**	< 0.001	0.577	0.736	**0.712**	< 0.001	0.595	0.852
	Quin	**0.656**	< 0.001	0.570	0.754	**0.749**	0.006	0.609	0.922
	Spl	**0.663**	< 0.001	0.583	0.755	**0.706**	< 0.001	0.583	0.853
RC									
	Lin	**0.651**	< 0.001	0.555	0.764	**0.647**	0.004	0.481	0.870
	Quin	**0.660**	< 0.001	0.545	0.800	**0.761**	0.091	0.555	1.044
	Spl	**0.665**	< 0.001	0.548	0.779	**0.676**	0.046	0.500	0.993
FFQ									
	Lin	**0.731**	< 0.001	0.650	0.822	**0.832**	0.053	0.691	1.002
	Quin	**0.723**	< 0.001	0.630	0.830	**0.824**	0.070	0.669	1.016
	Spl	**0.734**	< 0.001	0.644	0.836	**0.899**	0.378	0.709	1.140
MCML									
	Lin	**0.654**	< 0.001	0.558	0.767	**0.667**	0.006	0.499	0.890
	Quin	**0.605**	< 0.001	0.462	0.792	**0.728**	0.281	0.408	1.296
	Spl	**0.669**	< 0.001	0.561	0.796	**0.710**	0.051	0.504	1.002

Table 3

Simulation results of logistic regression for 500 simulated data sets. Displayed are the mean relative risks of moving from the 10th to the 90th percentile of the HEI-2005 total score in the linear analysis and from the 1st to the 5th quintile in the quintile analysis (Rel. Risk) across the simulation, 10 × the standard deviation across the simulations (sd). “Linear” risk function means the disease status is simulated from a logistic model in which the predictor total score enters linearly. “Quintile” risk function means the disease status is simulated from a logistic model which contains the dummy variables of the total score based on quintiles of the total score as predictors. The fit function is “Linear” if the total score enters the model linearly and is “Quintile” if we compare the relative risk of the 1st and 5th quintiles when fitting the model. The methods used are moment reconstruction (MR), moment-adjusted imputation (MAI), Monte Carlo maximum likelihood (MCML), and regression calibration (RC).

			Men				
Risk Function	Fit Function		Truth	MR	MAI	MCML	RC
Linear	Linear	RR	0.682	0.681	0.677	0.685	0.685
		10×sd		0.502	0.4810	0.474	0.474
Linear	Quintile	RR	0.682	0.661	0.654	0.635	0.662
		10×sd		0.574	0.554	0.592	0.559
Quintile	Quintile	RR	0.691	0.711	0.704	0.688	0.696
		10×sd		0.624	0.602	0.651	0.599
			Women				
Risk Function	Fit Function		Truth	MR	MAI	MCML	RC
Linear	Linear	RR	0.703	0.711	0.707	0.716	0.715
		10×sd		1.058	0.977	1.000	1.000
Linear	Quintile	RR	0.703	0.691	0.680	0.659	0.688
		10×sd		1.273	1.171	1.273	1.206
Quintile	Quintile	RR	0.712	0.748	0.741	0.708	0.715
		10×sd		0.913	0.857	0.891	0.819