HHS Public Access

Author manuscript

Biom J. Author manuscript; available in PMC 2017 November 01.

Published in final edited form as:

Biom J. 2016 November ; 58(6): 1538–1551. doi:10.1002/bimj.201500201.

The impact of stratification by implausible energy reporting status on estimates of diet-health relationships

Janet A. Tooze[*,1], **Laurence S. Freedman**[2], **Raymond J. Carroll**[3], **Douglas Midthune**[4], and **Victor Kipnis**[4]

[1]Department of Biostatistical Sciences, Wake Forest School of Medicine, Medical Center Blvd, Winston-Salem, NC 27157, USA

[2]Sheba Medical Center, Gertner Institute for Epidemiology and Health Policy Research, Tel Hashomer 52161, Israel

[3]Department of Statistics, Texas A&M University, College Station, TX 77843, USA

[4]National Cancer Institute, Bethesda, MD 20892, USA

Abstract

The food frequency questionnaire (FFQ) is known to be prone to measurement error. Researchers have suggested excluding implausible energy reporters (IERs) of FFQ total energy when examining the relationship between a health outcome and FFQ-reported intake to obtain less biased estimates of the effect of the error-prone measure of exposure; however, the statistical properties of stratifying by IER status have not been studied. Under certain assumptions, including nondifferential error, we show that when stratifying by IER status, the attenuation of the estimated relative risk in the stratified models will be either greater or less in both strata (implausible and plausible reporters) than for the nonstratified model, contrary to the common belief that the attenuation will be less among plausible reporters and greater among IERs. Whether there is more or less attenuation depends on the pairwise correlations between true exposure, observed exposure, and the stratification variable. Thus exclusion of IERs is inadvisable but stratification by IER status can sometimes help. We also address the case of differential error. Examples from the Observing Protein and Energy Nutrition Study and simulations illustrate these results.

Keywords

Attenuation; Bias (epidemiology); Food frequency questionnaire; Models; Statistical; Underreporting

[*]Corresponding author: jtooze@wakehealth.edu, Phone: +1-336-716-3833, Fax: +1-336-716-6427.

Additional supporting information including source code to reproduce the results may be found in the online version of this article at the publisher's web-site

Conflict of Interest

The authors have declared no conflict of interest.

1 Introduction

It is well established that self-report dietary assessments, such as the food frequency questionnaire (FFQ), and 24-hour recall are prone to both random and systematic measurement error (Thompson and Subar, 2013; Freedman et al., 2014). Researchers who study diet and health outcomes have developed statistical methods to describe and quantify the error in self-report instruments. One early and widely used method of identifying individuals who report their diets with error is the Goldberg method (Goldberg et al., 1991; Black et al., 1991). This method uses the ratio of reported energy intake to predicted energy needs from an equation to classify "implausible energy reporters" (IER), that is those persons that report energy intake that is either too high, termed over-reporters (OR), or too low, termed underreporters (UR), based on their predicted energy needs. Studies that have used this method and extensions thereof (Huang et al., 2004; Huang et al., 2005) have demonstrated that a large proportion of adults and children report implausible energy intake on self-report assessments, predominantly in the direction of UR (Macdiarmid and Blundell, 1998; Hill and Davies, 2001; Forrestal, 2011).

Some authors have suggested excluding IERs (Drummond et al., 1998; McCrory et al., 2002; Huang et al., 2004; Huang et al., 2005) from studies of diet and health to obtain more valid conclusions. This work has been motivated by the observation that when IERs are excluded from regression analyses of body mass index (BMI) on dietary intake, regression coefficients of the dietary intakes in the sample of plausible or acceptable reporters (AR) have been found stronger than the estimates from the total sample, and in some cases the regression coefficients change direction in the AR sample (Drummond et al., 1998; Huang et al., 2004; Howarth et al., 2005; Savage et al., 2008; Mendez et al., 2011; Noel et al., 2011). Recently, Rhee et al. (2015) have challenged the use of weight-based definition of IER status for this purpose, and recommend using plausible fixed limits for energy intake, such as whether the reported energy intake falls within plausible fixed limits regardless of the individual's body weight, such as a lower limit of 500 kcal/d and an upper limit of 3500 kcal/d (Willet, 1998). Although, our enquiry in this paper does not focus on the question of excluding IERs from the analysis, the material we will present does have implications for this question, which we will mention in the Discussion section.

Rather than excluding IERs from analyses, other authors have recommended stratification of diet-health analyses by IER status (Börnhorst et al., 2013) or using indicator variables to identify misreporters (Nielsen and Adair, 2007; Mendez et al., 2011). In this paper, we focus on the strategy of stratification by IER status using the Goldberg method, and examine the statistical properties of such stratification. We treat the cases of nondifferential error, where the outcome variable is independent of the error in the dietary report (Flegal et al., 1991; Carroll et al., 2006), and differential error separately. The nondifferential error assumption is often reasonable in nutritional cohort studies where the outcome is a future disease event and individuals do not report dietary intake differently based in a manner that is associated with a health outcome. In such studies, nondifferential error usually attenuates the estimate of the risk parameter (Freedman et al., 2011); the attenuation factor is a multiplicative factor that quantifies this underestimation. We show here that under nondifferential error, when stratifying by IER status, the attenuation factors within the UR stratum and the AR stratum

will both be either larger or smaller than for the nonstratified model. Whether there is an increase or decrease in the attenuation factor depends on a simple function of the pairwise correlations between true exposure (an estimate that is obtained through use of an unbiased biomarker), observed exposure, and the stratification variable. Since larger attenuation factors lead to less bias in the estimated risk parameter, this simple function provides a guide as to when risk estimates will be improved by stratification by IER status, and when they will not.

In studies where the outcome is body mass index (BMI), which is commonly associated with misreporting (Macdiarmid and Blundell, 1998), or whenever the study has a retrospective case-control design, a design prone to recall bias, the measurement error is likely to be differential. We address the case of differential error and show that stratification by IER status may increase or decrease the bias in the estimated risk parameter dependent on the correlations between the outcome, observed exposure, and the stratification variable.

We illustrate our results using data on protein and protein density (protein divided by energy) intakes from the Observing Protein and Energy Nutrition (OPEN) Study, which measured urinary nitrogen as a biomarker for protein intake and doubly labeled water for energy intake, and collected self-reported protein and energy intakes from a food frequency questionnaire. We also provide simulations to illustrate the cases of nondifferential and differential error.

2 Methods

2.1 Nondifferential error

We denote by Y the health status outcome. After transformation to a suitable scale (see below), true dietary intake of a nutrient of interest is denoted by X; the error-prone reported intake of this nutrient by W; and the ratio of reported energy intake from an instrument such as an FFQ to estimated basal metabolic rate is denoted by G. This represents the use of the Goldberg method to determine IER status, by which we stratify.

We assume that we can find suitable transformations (such as power or logarithmic) so that (X, W, G) are multivariate normal with means (μ_X, μ_W, μ_G), standard deviations (σ_X, σ_W, σ_G), and correlations (ρ_{XW}, ρ_{XG}, ρ_{WG}). Note that by assuming the multivariate normal distribution, we postulate that there is no interaction between G and W in the regression of X on these two variables. We assume a generalized linear model for the relationship between Y and X, where Y may be expressed in terms of a binary, continuous, or time to event variable:

$$\mathrm{E}(Y|X) = H(\alpha_0 + \alpha_X X) \quad (1)$$

Results apply to any model with a linear predictor. The primary parameter of interest in this model is α_X, which describes the relationship between Y and the true dietary intake, X. However, because X is unknown, W is substituted for X. For the overall (nonstratified) model, it has been shown that, under the nondifferential error assumption, the regression coefficient for this model is approximately equal to the regression coefficient for X

multiplied by an attenuation factor, which we call λ_{XW} (Kipnis et al., 2001). As in other studies (Kipnis et al., 2001; Kipnis et al., 2003), we assume the model for W for an individual i is related to X by:

$$W_i = \beta_{W0} + \beta_{W1} X_i + \delta_i, \quad (2)$$

where the FFQ is assumed to have systematic error (B_{W0} and B_{W1}, respectively) as well as random error, δ_i with mean 0 and variance σ^2_δ. We can estimate σ^2_X as the between subject variance of the unbiased biomarker of the nutrient M measured on individual i at time j on a transformed scale assuming the mixed model:

$$M_{ij} = X_i + u_{ij} \quad (3)$$

where u_{ij} is the within-person random error with variance σ^2_u. The attenuation factor λ_{XW} is given by:

$$\lambda_{XW} = \frac{\text{cov}\ (X_i, W_i)}{\sigma^2_W} = \frac{\beta_{W1}}{\sigma^2_W / \sigma^2_X} = \frac{\beta_{W1}}{\beta^2_{W1} + \sigma^2_\delta / \sigma^2_X}. \quad (4)$$

We now consider the stratified model. We consider two strata (ARs and URs), excluding ORs who comprise usually a very small group. The strata are defined according to whether or not $G \in A$, $A = (g_l, g_u)$. We assume that the error in W is also nondifferential within the strata defined by G. This occurs as long as the distribution of Y on X and G is the same as that of Y on X, that is G conveys no information about Y over and above that in X. This seems reasonable, since G is a variable that measures underreporting, and the assumption dovetails with the assumption of nondifferential error in W. In the Appendix, under these assumptions, we derive that the attenuation factor within any stratum A is:

$$\lambda_{XW|G \in A} = \lambda_{XW} \left[\frac{1 - \frac{\rho_{XG}\rho_{WG}}{\rho_{XW}} f_A}{1 - \rho^2_{WG} f_A} \right] \quad (5)$$

where ρ_{XG}, ρ_{WG}, and ρ_{XW} are the correlations between X and G, W and G, and X and W, respectively, and f_A is a factor dependent on the set A, but always lying between 0 and 1. The ratio in the brackets in Eq. (5) will therefore be greater than 1 and the attenuation factors for the stratified model will be greater than the attenuation factor for the full model if and only if

$$\frac{\rho_{XG}\rho_{WG}}{\rho_{XW}} < \rho^2_{WG} \quad (6)$$

The correlation between W and G may be determined empirically. Correlations with X may be calculated when an unbiased, or recovery biomarker, for X is available; they may also sometimes be estimated from calibration studies that employ a concentration biomarker that has been previously validated in a feeding study (see e.g. Freedman et al., 2010).

2.2 Differential error

Relating dietary intakes to outcomes when the dietary measurement error is differential is problematic. Unlike in the case of nondifferential error where the estimated risk parameter is (nearly) always attenuated, with the differential error there can be underestimation or overestimation of the risk parameter. (Sometimes, when the measurement error is well understood, the direction of bias in the estimated risk parameter can be predicted. See for example, our comments in the Discussion section regarding the case where body mass index (BMI) is the outcome measure.) Likewise, the effect of stratification by IER status is difficult to predict. Firstly, if the direction of the bias cannot be anticipated, it is unknown whether it is better to attenuate or de-attenuate the estimated risk parameter. Secondly, in the case of differential error, Eq. (5) does not hold, and consequently condition (6) does not guarantee an increase in the attenuation factor. One can show for the linear regression of Y on W and of Y on W and G, that if the coefficient of W in the model with W and G and the coefficient in the model with just W have the same sign, then the absolute value of the coefficient in the model with W and G will be larger than that in the model with W, only if

$$\frac{\rho_{YG}\rho_{WG}}{\rho_{YW}} < \rho_{WG}^2 \quad (7)$$

(see Appendix for proof). This condition looks similar to that in (6), but Y now takes the place of X, making it difficult to predict in advance the effect of stratification by IER status. Also, unlike condition (6), condition (7) only applies when there is no change of sign in the coefficient; however, such a condition cannot be guaranteed. In the Results section, we present simulations of differential error, in a case where there is no true association between diet and outcome, so that any nonnull estimate is biased, and it will be advantageous if we can induce further attenuation of the risk estimate by stratification. We will show that stratification by IER status can sometimes increase and sometimes decrease the bias in the estimate of the coefficient. Note that for nondifferential error, Eq. (7) also holds, in addition to Eq. (6).

2.3 OPEN study methods

The OPEN Study was conducted to estimate the measurement error properties of self-report dietary assessment including the Diet History Questionnaire, a food frequency questionnaire (FFQ) for dietary assessment (Subar et al., 2003). Participants were recruited from a random sample of 5000 households in the Washington, DC, metro area with a household member aged 40–69 years; the final sample size was 484. Participants completed the FFQ twice, approximately 3 months apart. In the interim, doubly labeled water was used to estimate total energy expenditure (and hence energy intake under stable weight conditions) over a two-week period, and urinary nitrogen, a recovery biomarker for protein intake, was

estimated twice from two 24-hour urine collections. The within-person variation in doubly labeled water was estimated by dosing 25 participants a second time approximately 2 weeks after the first data collection. Weight was measured under standard conditions at all visits, and height was measured at visit 1. Basal metabolic rate was calculated from height, weight, and age at visit 1 using the equation developed by Schofield (1985). The Goldberg method (Black et al., 1991; Goldberg et al., 1991) was used to classify participants as UR, OR, or AR. We used values for the coefficient of variation suggested by Black (2000) to classify IER status. All analyses were done in SAS (version 9.4, Cary, NC).

3 Results

3.1 Example: OPEN Study

The final sample size of participants with valid doubly labeled water measurements, at least 1 valid urinary nitrogen measure, and the food frequency questionnaire in the OPEN Study was 388. Using the cutpoints suggested by Black (2000; 1.10 and 2.19), 51% of women and 52% of men were classified as UR on the food frequency questionnaire; only 1.9% of women and 2.5% of men were classified as OR. ORs were excluded from the analysis (N = 13, resulting in a final analytic sample of N = 375 comprising 200 men and 175 women). The estimated means, variances, and pairwise correlations of log FFQ reported intake (W), the log recovery biomarker (M), the latent variable log true intake (X) (inferred from M), and the log ratio of reported energy intake to estimated basal metabolic rate (G) for men are presented in Table 1, both for absolute protein intake and for protein density intake. We used Eq. (4) to estimate the attenuation factor, overall and within the AR and UR strata for absolute protein and for protein density as assessed by the FFQ; the standard error was estimated using the delta method.

Results (Table 2) indicated a substantial improvement in the attenuation factor in the models stratified by AR and UR status for the absolute protein model (from 0.168 in the overall model to 0.381 for ARs and 0.288 for URs), consistent with the ratio predicted from Eq. (5). However, for protein density, the attenuation factors were comparable between the overall model and the stratified models, a result that was also consistent with the Eq. (5) prediction. As seen in Table 1, although estimated correlations between X and W were quite similar for protein and protein density and correlations between X and G were both small, correlations between W and G differed greatly for absolute protein and protein density, leading to these disparate results.

3.2 Validation of the method: Simulation studies

To check our methods, we performed a simulation study to examine the impact of stratifying by underreporting status in epidemiologic studies under the assumption of nondifferential error. First, we generated a multivariate normal distribution of height, weight, age, log FFQ energy (from first FFQ), log FFQ protein intake (from first FFQ), log true energy intake, and log true protein intake for men, using parameters estimated from the data of the OPEN Study. We then simulated 1200 datasets from this distribution for a hypothetical nested case-control study with 1000 cases and 1000 controls. We classified participants by IER status using the standard Goldberg cutoffs (Black, 2000), and excluded ORs, leaving URs and

ARs. We simulated the outcome Y as a function of true protein intake X_P or true protein density X_{PD}:

$$\text{logit}\ (\text{P}\,(Y=1)) = -\,7.852+1.30X_{P},$$

and

$$\text{logit}\ (\text{P}\,(Y=1)) = 3.8516+2.0X_{PD}$$

We regressed true protein (protein density) on the FFQ value to estimate attenuation and correlation between truth and the FFQ, and used logistic regression to estimate α_{X_P} and $\alpha_{X_{PD}}$ overall and in stratified models. Results (Table 3) confirmed that for absolute protein the estimated log odds ratios were higher within the AR and UR strata than overall, whereas for protein density they were similar. Also note that for absolute protein, the proportion of bootstrap samples in which the null hypothesis was rejected is similar for the overall model to each of the strata, but in a stratified model (including an indicator variable for the strata), this proportion is increased. For protein density, there is a power loss by restricting the analysis to just one strata, and little change in the stratified model.

We conducted a simulation study to investigate the impact of differential error on stratification by underreporting status when the true relationship with the outcome and true intake was null. Our particular interest was whether spurious relationships between dietary intake and outcome would be produced by such stratification. We generated a multivariate normal distribution of log BMI (Y), height, age, log FFQ energy intake, log FFQ protein intake (W, from first FFQ), and log protein (X) (estimated from urinary nitrogen), using parameters estimated from the data of the OPEN Study (see Table 1). We then set the correlation between Y and X to zero to simulate a null relationship, that is $Y = \alpha_0 + e_Y$, and we also varied cov(Y, W) in the simulations. We simulated 1200 datasets from this distribution for a hypothetical study with 2000 participants. We classified participants by misreporting status using the standard Goldberg cutoffs (Black, 2000) (grouping OR with AR) and then regressed Y on W. Results are shown in Table 4. It can be seen that, as expected, the differential error in the measurement of dietary intake W, creates a spuriously nonzero estimate of the overall slope. Stratification by IER status sometimes decreases this bias and sometimes increases it, depending on the correlation structure, and according to the criterion given in Eq. (7).

4 Discussion

Under nondifferential error, we show that when stratifying by IER status, the attenuation factors within the UR stratum and the AR stratum will both be either greater or smaller than for the nonstratified model. In the OPEN study stratifying by underreporting status resulted in larger attenuation factors than in the overall group for protein expressed in kcal/day, with relative increases in attenuation factors by 128% for men and 83% for women. This was largely driven by the strong correlation between the protein intake reported on the FFQ and the ratio of reported energy intake to predicted total energy expenditure (ρ_{WG}, 0.80 for men

and 0.82 for women). However, when protein was expressed as a density (% of energy from protein), this correlation dropped to −0.09 for men and −0.03 for women, resulting in a ratio of approximately one, and, therefore, not leading to a change in the attenuation factors with stratification by misreporting status. The gains in power for analysis of protein and a health outcome that may be obtained from stratifying by IER status, and the gains in power obtained by studying protein density rather than protein both appear to derive from the same source, namely, the correlation between reported protein intake and reported energy intake. It is principally this correlation that leads to the high value for ρ_{WG}, and it is also this correlation that renders FFQ-reported protein density to have a higher correlation with true intake than does reported absolute protein intake. Thus, one may speculate that gains from stratification by IER status will be found more in analyses of absolute intakes than in analyses of densities.

Our analyses strongly suggest that excluding IERs from the analysis is not an optimal strategy, contrary to previous suggestions (Drummond et al., 1998; McCrory et al., 2002; Huang et al., 2004; Huang et al., 2005). When it appears beneficial (i.e., when criterion (6) is satisfied), to stratify by IER status, then the benefit is found in both groups, so excluding one of them would unnecessarily lose statistical power relative to a stratified analysis. When stratification by IER status is not beneficial, then exclusion of IERs is equivalent to stratification followed by discarding one of the strata, and would be inferior to the unstratified analysis.

The statistical model from which we derive these results does not include any interaction between the IER stratification variable and reported intake in the regression of true intake on these two variables. Indeed, the multivariate normal distribution assumed by us precludes such an interaction. However, the argument made by those who advocate excluding URs is that they dilute the relationship between a dietary intake and the outcome of interest, which implies a considerably lower attenuation factor among URs than among ARs that would occur only if there were a large interaction between reported intake and IER status. To date, there is little indication that such large interactions are to be found. There has been some examination of interactions between reported intake and personal characteristics in calibration equations that relate reported intake to true intake. For FFQ reports of energy and protein intake, Neuhouser et al. (2008) reported some interaction with BMI, as did Freedman et al. (2014) among women but not men, but these were moderate in size. We are unaware of any investigation of interactions with IER status as defined by the Goldberg method. We therefore examined this in the OPEN Study data and did not find evidence of such interaction. We performed, separately for men and women, linear regression of log urinary protein on log FFQ-reported protein, underreporter status (UR or AR) and their interaction term. The interaction term was modest and statistically nonsignificant for both genders.

In all of the papers, we identified on accounting for IERs in studying diet and health outcome relationships (Drummond et al., 1998; McCrory et al., 2002; Huang et al., 2004; Howarth et al., 2005; Huang et al., 2005; Fiorito et al., 2006; Nielsen and Adair, 2007; Savage et al., 2008; Mendez et al., 2011; Noel et al., 2011; Börnhorst et al., 2013), all of them modeled BMI or adiposity as the outcome of interest. This is problematic due to the

nondifferential error expected, since overweight or obese persons are known to underreport their energy intake on average to a greater extent than other persons. This relationship between overweight and underreporting will tend to turn a true-positive association into an estimated one that is even more attenuated than under nondifferential error, or even into an estimated negative association. Even though one could predict the direction of the bias due to differential reporting error, stratifying by IER status could either reduce or magnify the bias in an unpredictable manner. Therefore, when BMI is the outcome of interest, it is not advised to use directly a FFQ or another instrument prone to nondifferential error to model associations, but to use an unbiased marker of the nutrient of intake, or the calibration equation strategy advocated by Prentice et al. (2011). If an error-prone instrument is used, the analytic assumptions used and their impact on limiting the interpretation of the results should be clearly addressed. Although we focused on BMI and differential error in this paper, certainly other variables may also be associated with both measurement error in a predictor variable and with the outcome and such a possibility should be carefully examined. When this occurs, the differential aspect of the error can be eliminated or reduced by including such variables in the disease model. Differential error should also be carefully considered when dietary data are collected retrospectively.

One limitation of this work is that our model assumes that just one dietary exposure is of interest. In reality, researchers are often interested in measuring two or more dietary or other exposures that may be measured with error. This may lead to residual confounding of the estimated effect of the dietary exposures of interest (Freedman et al., 2011), in which case these results may not hold. However, for dietary protein, the amount of residual confounding has been estimated to be small (Freedman et al., 2011).

In summary, it may sometimes be beneficial to stratify by underreporting status (or include it as an indicator variable) when modeling diet and disease relationships, and we present the conditions under which incorporating misreporting status into the analysis results in more favorable attenuation factors. However, this strategy is safely employed only when measurement error is nondifferential, and can lead to further bias when dietary measurement error is differential, as in the case when BMI is the outcome variable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Cancer Institute at the National Institutes of Health (P30 CA012197 for J.A.T. and U01 CA057030 for R.J.C.).

Appendix

A.1 Proof that the error in *W* is nondifferential within the strata defined by *G*

We assume that W has nondifferential error, that is

$$f(Y|X,W) = f(Y|X) \quad \text{(A1)}$$

and we also assume G provides no information on Y conditional on X. This means that

$$f(Y|X,G) = f(Y|X) \quad \text{(A2)}$$

We can show that W has non-differential error within the strata of G, that is $f(Y|X, G, W) = f(Y|X, G)$. The proof of this when (Y, X, G, W) are multivariate normal and therefore linearly related, then the coefficient for W in the regression of Y on X, G, and W is proportional to:

$$\sigma_{YX}\left(\sigma_{XG}\sigma_{GW} - \sigma_{XW}\sigma_G^2\right) + \sigma_{YW}\left(\sigma_X^2\sigma_G^2 - \sigma_{XG}^2\right) + \sigma_{YG}\left(\sigma_{XG}\sigma_{XW} - \sigma_X^2\sigma_{GW}\right) \quad \text{(A3)}$$

From (A1), $\sigma_{YW} = \sigma_{YX}\sigma_{XW}/\sigma_X^2$ and from (A2) $\sigma_{YG} = \sigma_{YX}\sigma_{XG}/\sigma_X^2$. Substituting these into (A3) we get that the coefficient of W equals zero. Therefore Y is independent of W conditional on X and G.

A.2 Derivation of attenuation factor for stratified model under assumption of nondifferential error

Let $A = [a_L, a_H]$, and suppose that (X, W, G) are multivariate normal with means (μ_X, μ_W, μ_G), standard deviations $(\sigma_X, \sigma_W, \sigma_G)$, and correlations $(\rho_{XW}, \rho_{XG}, \rho_{WG})$. The attenuation in the unstratified data is $\lambda_{XW} = \text{cov}(X, W)/\text{var}(W)$. We aim to compute the attenuation for stratified data when $G \in A$, i.e. $\lambda_{XW|G\in A} = \text{cov}(X, W|G \in A)/\text{var}(W|G \in A)$, and to show when it is that the attenuation coefficient the full data model is smaller than the attenuation coefficient for the stratified data.

We have the following result:

Lemma 1

The attenuation in the stratified model will be greater than that of the full model if $\rho_{XG}\rho_{WG}/\rho_{XW} < \rho_{WG}^2$.

Here is the technical argument.

Define $X_* = (X - \mu_X)/\sigma_X$ and similarly for W_* and G_*, thus standardizing all three. Also, define the standardized interval $A_* = \{a_{*L} = (a_L - \mu_G)/\sigma_G, a_{*H} = (a_H - \mu_G)/\sigma_G\}$. It is clear that

$$\begin{aligned}\text{cov}\ (X, W|G \in A) &= \sigma_X\sigma_W \text{cov}\ (X_*, W_*|G_* \in A_*)\,;\\ \text{var}\ (W|G \in A) &= \sigma_W^2 \text{var}\ (W_*|G_* \in A_*)\,,\end{aligned}$$

and hence the attenuation coefficient for the stratified sample when $G \in A$ is

$$\lambda_{XW|G\in A} = (\sigma_X/\sigma_W)\,\mathrm{cov}\,(X_*, W_*|G_* \in A_*)\,/\mathrm{var}\,(W_*|G_* \in A_*)\,. \quad \text{(A4)}$$

The random variable G_* given $G_* \in A_*$ is a standard normal random variable truncated to the set A_*, and consequently its density is $\phi(g_*)/\{\Phi\,(a_{*H}) - \Phi\,(a_{*L})\}$, where $\phi(\cdot)$ is the standard normal density function and $\Phi(\cdot)$ is the corresponding cumulative distribution function. Define

$$\theta = \{\phi\,(a_{*L}) - \phi\,(a_{*H})\}\,/\,\{\Phi\,(a_{*H}) - \Phi\,(a_{*L})\}$$
$$\eta = \{a_{*L}\phi\,(a_{*L}) - a_{*H}\phi\,(a_{*H})\}\,/\,\{\Phi\,(a_{*H}) - \Phi\,(a_{*L})\}\,.$$

It is known from properties of the truncated normal distribution that $E(G_*|G_* \in A_*) = \theta$ and $\mathrm{var}(G_*|G_* \in A_*) = 1 + \eta - \theta^2$. This means that $\mathrm{E}(W_*|G_*) = \rho_{WG}G_*$, and consequently

$$\mathrm{E}\,(W_*|G_* \in A_*) = \mathrm{E}\,\{\mathrm{E}\,(W_*|G, G_* \in A_*)\,|G_* \in A_*\} = \mathrm{E}\,(\rho_{WG}G_*|G_* \in A_*) = \rho_{WG}\theta.$$

Since $\mathrm{var}\,(W_*|G_*) = E\left(W_*^2|G_*\right) = 1 - \rho_{WG}^2$,

$$\begin{aligned}\mathrm{E}\,(W_*^2|G_* \in A_*) &= \mathrm{E}\,\{\mathrm{E}\,(W_*^2|G_*)\,|G_* \in A_*\} = \mathrm{var}\,(W_*|G_*) + \rho_{WG}^2\,\mathrm{E}\,(G_*^2|G_* \in A_*) = \\ &= \mathrm{var}\,(W_*|G_*) + \rho_{WG}^2\,\mathrm{var}\,(G_*|G_* \in A_*) - \rho_{WG}^2\{\mathrm{E}\,(G_*|G_* \in A_*)\}^2 = \\ &= 1 - \rho_{WG}^2 + \rho_{WG}^2\,(1+\eta-\theta^2) + \rho_{WG}^2\theta^2.\end{aligned}$$

Consequently, since $\mathrm{E}(W_*|G_* \in A_*) = \rho_{WG}\theta$,

$$\mathrm{var}\,(W_*|G_* \in A_*) = 1 - \rho_{WG}^2 + \rho_{WG}^2\left(1+\eta-\theta^2\right) = 1 + \rho_{WG}^2\left(\eta-\theta^2\right). \quad \text{(A5)}$$

Similarly, since $\mathrm{cov}(X_*,\ W_*|G_*) = \rho_{XW} - \rho_{XG}\rho_{WG}$, $\mathrm{E}(X_*|G_*) = \rho_{XG}G_*$ and $\mathrm{E}(X_*|G_* \in A_*) = \rho_{XG}\theta$,

$$\begin{aligned}\mathrm{E}\,(X_*W_*|G_* \in A_*) &= \mathrm{cov}\,(X_*, W_*|G_*) + \rho_{WG}\rho_{XG}\mathrm{E}\,(G_*^2|G_* \in A_*) = \\ &= \rho_{XW} - \rho_{XG}\rho_{WG} + \rho_{WG}\rho_{XG}\,\mathrm{var}\,(G_*|G_* \in A_*) + \\ &\quad -\rho_{WG}\rho_{XG}\{\mathrm{E}\,(G_*|G_* \in A_*)\}^2 = \\ &= \rho_{XW} - \rho_{XG}\rho_{WG} + \rho_{WG}\rho_{XG}\,(1+\eta-\theta^2) - \rho_{WG}\rho_{XG}\theta^2,\end{aligned}$$

so that

$$\mathrm{cov}\,(X_*, W_*|G_* \in A_*) = \rho_{XW} + \rho_{WG}\rho_{XG}\left(\eta-\theta^2\right). \quad \text{(A6)}$$

Combining (A4), (A5), and (A6), and remembering that without stratification the attenuation is $\lambda_{XW} = \rho_{XW}\,\sigma_X/\sigma_W$, we have

$$\lambda_{XW|G\in A}=\frac{\sigma_X}{\sigma_W}\frac{\rho_{XW}+\rho_{WG}\rho_{XG}\,(\eta-\theta^2)}{1+\rho_{WG}^2\,(\eta-\theta^2)}=\lambda_{XW}\frac{1+\rho_{WG}\rho_{XG}\,(\eta-\theta^2)\,/\rho_{XW}}{1+\rho_{WG}^2\,(\eta-\theta^2)}. \quad \text{(A7)}$$

Since var($G_*|G_* \in A_*$) = 1 + $\eta - \theta^2$, it is necessarily the case that $-1 \le \eta - \theta^2 \le 0$. Substitution of $f_A = \theta^2 - \eta$ yields Eq. (5). Consequently, the attenuation in the stratified model will be greater than that of the full model if $\rho_{XG}\rho_{WG}/\rho_{XW} < \rho_{WG}^2$.

A.3 Derivation of condition (6) for normally distributed outcomes

In the linear regression of Y on W, the coefficient of W is given by $\beta_W = \sigma_{YW}/\sigma_W^2$. In the linear regression of Y on W and G, the coefficient of W is given by

$$\beta_{W|G} = \left(\sigma_{YW}\sigma_G^2 - \sigma_{YG}\sigma_{WG}\right) / \left(\sigma_W^2\sigma_G^2 - \sigma_{WG}^2\right).$$

Suppose that $\sigma_{YW} > 0$. Then $\beta_W < \beta_{W|G}$ implies that $\sigma_{YG}\sigma_{WG}\sigma_W^2 < \sigma_{YW}\sigma_{WG}^2$, which, after substituting correlation for covariance, $\sigma_{YG} = \rho_{YG}\sigma_Y\sigma_G$, and some algebraic manipulation, leads to condition (6) in the text. Assuming $\sigma_{YW} < 0$, the condition $\beta_W > \beta_{W|G}$ also leads to condition (6). Note that we assume in this proof that β_W and $\beta_{W|G}$ have the same sign.

References

Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. European Journal of Clinical Nutrition. 1991; 45:583–599. [PubMed: 1810720]

Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake, basal metabolic rate. A practical guide to its calculation, use and limitations. International Journal of Obesity and Related Metabolic Disorders. 2000; 24:1119–1130. [PubMed: 11033980]

Börnhorst C, Huybrechts I, Hebestreit A, Vanaelst B, Molnár D, Bel-Serrat S, Mouratidou T, Moreno LA, Pala V, Eha M, Kourides YA, Siani A, Eiben G, Pigeot I, IDEFICS consortium. Diet-obesity associations in children: approaches to counteract attenuation caused by misreporting. Public Health Nutrition. 2013; 16:256–266. [PubMed: 23046605]

Carroll, RJ.; Ruppert, D.; Stefanski, LA.; Crainiceanu, C. Measurement Error in Nonlinear Models: A Modern Perspective. 2nd. CRC Press; Boca Raton, FL: 2006.

Drummond SE, Crombie NE, Cursiter MC, Kirk TR. Evidence that eating frequency is inversely related to body weight status in male, but not female, non-obese adults reporting valid dietary intakes. International Journal of Obesity and Related Metabolic Disorders. 1998; 22:105–112. [PubMed: 9504318]

Fiorito LM, Ventura AK, Mitchell DC, Smiciklas-Wright H, Birch LL. Girls' dairy intake, energy intake, and weight status. Journal of the American Dietetics Association. 2006; 106:1851–1855.

Flegal KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. American Journal of Epidemiology. 1991; 134:1233–1244. [PubMed: 1746532]

Forrestal SG. Energy intake misreporting among children and adolescents: a literature review. Maternal and Child Nutrition. 2011; 7:112–127.

Freedman LS, Kipnis V, Schatzkin A, Tasevska N, Potischman N. Can we use biomarkers in combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? Epidemiologic Perspectectives and Innovations. 2010; 7:2. Online manuscript, 9 pages.

Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. Journal of the National Cancer Institute. 2011; 103:1086–1092. [PubMed: 21653922]

Freedman LS, Commins JM, Moler JE, Willett W, Tinker LF, Subar AF, Spiegelman D, Rhodes D, Potischman N, Neuhouser ML, Moshfegh AJ, Kipnis V, Arab L, Prentice RL. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. American Journal of Epidemiology. 2014; 180:172–188. [PubMed: 24918187]

Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. European Journal of Clinical Nutrition. 1991; 45:569–581. [PubMed: 1810719]

Hill RJ, Davies PS. The validity of self-reported energy intake as determined using the doubly labelled water technique. British Journal of Nutrition. 2001; 85:415–430. [PubMed: 11348556]

Howarth NC, Huang TT, Roberts SB, McCrory MA. Dietary fiber and fat are associated with excess weight in young and middle-aged US adults. Journal of the American Dietetics Association. 2005; 105:1365–1372.

Huang TT, Howarth NC, Lin BH, Roberts SB, McCrory MA. Energy intake and meal portions: associations with BMI Percentile in U.S. Children. Obesity Research. 2004; 12:1875–1885. [PubMed: 15601985]

Huang TT, Roberts SB, Howarth NC, McCrory MA. Effect of screening out implausible energy intake reports on relationships between diet and BMI. Obesity Research. 2005; 13:1205–1217. [PubMed: 16076990]

Kipnis V, Midthune D, Freedman LS, Bingham S, Schatzkin A, Subar A, Carroll RJ. Empirical evidence of correlated biases in dietary assessment instruments and its implications. American Journal of Epidemiology. 2001; 153:394–403. [PubMed: 11207158]

Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, Bingham S, Schoeller DA, Schatzkin A, Carroll RJ. The structure of dietary measurement error: results of the OPEN biomarker study. American Journal of Epidemiology. 2003; 158:14–21. [PubMed: 12835281]

Macdiarmid J, Blundell J. Assessing dietary intake: who, what and why of under-reporting. Nutrition Research Reviews. 1998; 11:231–253. [PubMed: 19094249]

McCrory MA, Hajduk CL, Roberts SB. Procedures for screening out inaccurate reports of dietary energy intake. Public Health Nutrition. 2002; 5:873–882. [PubMed: 12633510]

Mendez MA, Popkin BM, Buckland G, Schroder H, Amiano P, Barricarte A, Huerta JM, Quirós JR, Sánchez MJ, González CA. Alternative methods of accounting for underreporting and overreporting when measuring dietary intake-obesity relations. American Journal of Epidemiology. 2011; 173:448–458. [PubMed: 21242302]

Nielsen SJ, Adair L. An alternative to dietary data exclusions. Journal of the American Dietetics Association. 2007; 107:792–799.

Noel SE, Ness AR, Northstone K, Emmett P, Newby PK. Milk intakes are not associated with percent body fat in children from ages 10 to 13 years. Journal of Nutrition. 2011; 141:2035–2041. [PubMed: 21940511]

Neuhouser ML, Tinker L, Shaw PA, Schoeller D, Bingham SA, Horn LV, Beresford SA, Caan B, Thomson C, Satterfield S, Kuller L, Heiss G, Smit E, Sarto G, Ockene J, Stefanick ML, Assaf A, Runswick S, Prentice RL. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. American Journal of Epidemiology. 2008; 167:1247–1259. [PubMed: 18344516]

Prentice RL, Mossavar-Rahmani Y, Huang Y, Van Horn L, Beresford SA, Caan B, Tinker L, Schoeller D, Bingham S, Eaton CB, Thomson C, Johnson KC, Ockene J, Sarto G, Heiss G, Neuhouser ML. Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. American Journal of Epidemiology. 2011; 175:591–603.

Rhee JJ, Sampson L, Cho E, Hughes MD, Hu FB, Willett WC. Comparison of methods to account for implausible reporting of energy intake in epidemiologic studies. American Journal of Epidemiology. 2015; 181:225–233. [PubMed: 25656533]

Savage JS, Mitchell DC, Smiciklas-Wright H, Symons Downs D, Birch LL. Plausible reports of energy intake may predict body mass index in pre-adolescent girls. Journal of the American Dietetics Association. 2008; 108:131–135.

Schofield WN. Predicting basal metabolic rate: new standards and review of previous work. Human Nutrition. Clinical Nutrition. 1985; 39:5S–41S.

Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, Sharbaugh CO, Trabulsi J, Runswick S, Ballard-Barbash R, Sunshine J, Schatzkin A. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. American Journal of Epidemiology. 2003; 158:1–13. [PubMed: 12835280]

Thompson, FE.; Subar, AF. Dietary assessment methodology. In: Coulston, AM.; Boushey, CJ.; Ferruzzi, MG., editors. Nutrition in the Prevention and Treatment of Disease. 3rd. Academic Press; San Diego, CA: 2013. p. 3-40.

Willett, WC. Nutritional Epidemiology. Oxford University Press; New York, NY: 1998.

Table 1

Distributions of W, X, and G in Men ($N = 200$) from the Observing Protein and Energy Nutrition Study, Washington, DC, Metro Area, 1999–2000.

Measure (log scale)	Mean	Variance	Correlation with X	Correlation with W
Absolute protein (log kcal)				
Protein, FFQ (W)	5.64	0.156	0.329	
Protein, UN (M)	6.04	0.063	0.800	0.263
Protein, truth (X)	6.04	0.040		0.329
rEI/BMR (G)	0.05	0.128	−0.060	0.803
Protein density				
% Energy protein, FFQ (W)	2.70	0.042	0.401	
% Energy protein, UN (M)	2.69	0.053	0.753	0.302
% Energy protein, truth (X)	2.69	0.030		0.401
rEI/BMR (G)	0.05	0.128	−0.025	−0.089
Other variables[a)]				
BMI (kg/m^2, log scale)	3.32	0.023	0.382	−0.004
Height (cm)	176.6	56.3	0.280	0.267
Age (y)	54.3	76.6	−0.054	−0.086

FFQ, food frequency questionnaire; UN, urinary nitrogen; rEI, reported energy intake from FFQ; BMR, basal metabolic rate.

[a)] Correlation with W for other variables is with absolute protein.

Table 2

Attenuation factors for stratified models in men ($N = 200$) from the Observing Protein and Energy Nutrition Study, Washington, DC, Metro Area, 1999–2000.

Attenuation factor (SE) (λ_{XW})	Strata	Lower cutoff(a_L)	Upper cutoff(a_L)	Lower cutoff (standardized) (a_{*L})	Upper cutoff (standardized) (a_{*U})	θ	η	Ratio	$\lambda_{XW\|\mathrm{Ge}\ A}$ predicted	$\lambda_{XW\|\mathrm{Ge}\ A}$ stratified (SE)
Absolute protein										
0.168 (0.040)	AR	0.095	0.784	0.120	2.042	0.802	−0.125	2.21	0.371	0.381 (0.076)
	UR	−∞	0.095	−∞	0.120	−0.723	−0.086	1.80	0.302	0.288 (0.063)
Protein density										
0.341 (0.070)	AR	0.095	0.784	0.120	2.042	0.802	−0.125	1.00	0.341	0.268 (0.103)
	UR	−∞	0.095	−∞	0.120	−0.723	−0.086	1.00	0.341	0.407 (0.096)

AR, acceptable reporters; UR, underreporters; for definitions of θ and η see Appendix.

Table 3

Parameter estimates and estimated attenuation factors from $N = 1200$ simulated datasets of size $N = 2000$.

	Average *N* cases	Average *N* controls	Intercept (SE)	Slope (SE)	Ratio (SE)	Expected attenuation (SE)	Attenuation (SE)	% slope Wald test $p < 0.05$
Absolute protein								
Simulated			−7.852	1.300				
Overall	979	985	−1.324 (0.668)	0.234 (0.118)			0.181 (0.011)	50.6%
AR strata	422	428	−2.770 (1.468)	0.466 (0.248)	2.05 (0.080)	0.371 (0.012)	0.362 (0.021)	47.3%
UR strata	557	557	−2.090 (1.125)	0.386 (0.207)	1.69 (0.051)	0.305 (0.012)	0.296 (0.018)	50.0%
Stratified	979	985	−2.36 (0.891)	0.417 (0.157)				76.3%
Protein density								
Simulated			−3.708	1.380				
Overall	981	983	−1.268 (0.611)	0.471 (0.225)			0.341 (0.018)	57.4%
AR	427	421	−1.232 (0.875)	0.454 (0.325)	0.994 (0.004)	0.339 (0.018)	0.338 (0.026)	24.5%
UR	553	562	−1.287 (0.839)	0.480 (0.308)	0.995 (0.003)	0.339 (0.018)	0.340 (0.024)	37.7%
Stratified	981	983	−1.263 (0.612)	0.469 (0.226)				56.9%

AR, acceptable reporters; UR, underreporters.

Table 4

Simulations under differential error when the true relationship with Y is null, N = 1200 simulated datasets of size N = 2000.

Predictor	r_{YW}	r_{YG}	r_{WG}	$\frac{r_{YG}r_{WG}}{r_{YW}}$	Expected relationship	Estimate of β_W or $\beta_{W\|G}$ (SE)	% $p < 0.05$
Overall	−0.400	−0.219	0.895	0.490	$\|\beta_W\| < \|\beta_{W\|G}\|$	−0.155 (0.008)	100%
UR						−0.211 (0.014)	100%
AR						−0.223 (0.017)	100%
Overall	−0.200	−0.219	0.850	0.931	$\|\beta_W\| > \|\beta_{W\|G}\|$	−0.077 (0.008)	100%
UR						−0.058 (0.015)	96.8%
AR						−0.060 (0.019)	92.8%
Overall	−0.004	−0.219	0.807	45.4	N/A	−0.002 (0.009)	4.4%
UR						0.068 (0.015)	99.7%
AR						0.073 (0.017)	98.9%
Overall	0.010	−0.219	0.803	−18.2	$\|\beta_W\| < \|\beta_{W\|G}\|$	0.004 (0.008)	5.8%
UR						0.076 (0.014)	99.9%
AR						0.081 (0.017)	99.7%
Overall	0.200	−0.219	0.761	−0.832	$\|\beta_W\| < \|\beta_{W\|G}\|$	0.077 (0.009)	100%
UR						0.182 (0.013)	100%
AR						0.190 (0.016)	100%

AR, acceptable reporters; UR, underreporters.