American Journal of Epidemiology

Vol. 170, No. 1
DOI: 10.1093/aje/kwp086
Advance Access publication April 30, 2009

Practice of Epidemiology

Comparing Methods for Accounting for Seasonal Variability in a Biomarker When Only a Single Sample Is Available: Insights From Simulations Based on Serum 25-Hydroxyvitamin D

Yiting Wang,* Eric J. Jacobs, Marjorie L. McCullough, Carmen Rodriguez,† Michael J. Thun, Eugenia E. Calle,† and W. Dana Flanders

Initially submitted October 22, 2008; accepted for publication March 13, 2009.

In biomarker-disease association studies, the long-term average level of a biomarker is often considered the optimal measure of exposure. Long-term average levels may not be accurately measured from a single sample, however, because of systematic temporal variation. For example, serum 25-hydroxyvitamin D (25(OH)D) concentrations may fluctuate because of seasonal variation in sun exposure. Association studies of 25(OH)D and cancer risk have used different strategies to minimize bias from such seasonal variation, including adjusting for date of sample collection (DOSC), often after matching on DOSC, and/or using season-specific cutpoints to assign subjects to exposure categories. To evaluate and understand the impact of such strategies on potential bias, the authors simulated a population in which 25(OH)D levels varied between individuals and by season, and disease risk was determined by long-term average 25(OH)D. Ignoring temporal variation resulted in bias toward the null. When cutpoints that did not account for DOSC were used, adjustment for DOSC sometimes resulted in bias away from the null. Using season- or month-specific cutpoints reduced bias toward the null and did not cause bias away from the null. To avoid potential bias away from the null, using season- or month-specific cutpoints may be preferable to adjusting for DOSC.

biological markers; epidemiologic measurement; seasons; vitamins

Abbreviations: DOSC, date of sample collection; 25(OH)D, serum 25-hydroxyvitamin D.

Measurement error is an important problem in epidemiologic studies of the association between a biomarker and disease (1). Disease risk may plausibly depend on an individual's long-term average level of a biomarker. Therefore, we assume that the disease risk is determined by the biomarker's long-term average level, referred to as the *true* level in this paper, and we consider it to be fixed for each individual. In some circumstances, temporal factors may influence short-term levels of the biomarker, so that a given single measure may not adequately represent long-term levels. For example, serum 25-hydroxyvitamin D (hereafter referred to as 25(OH)D) concentrations fluctuate because of seasonal variation in sun exposure (2). In large epidemiologic studies, often only one biomarker sample is available for each subject, and samples from different subjects are collected on different dates (e.g., in different months or seasons). If a single measured level is used to reflect the long-term average, measurement error is likely. For instance, subjects whose samples happened to be collected during the winter rather than the summer may be misclassified into low-level 25(OH)D categories.

Researchers have used different analytic strategies to address the problem of measurement error caused by temporal variation when only one sample is available per subject. In nested case-control studies of the metabolic precursor of vitamin D, some researchers matched on (and adjusted for) date of sample collection (DOSC) but did not account for DOSC when creating 25(OH)D exposure categories

* Correspondence to Dr. Yiting Wang, Department of Epidemiology and Surveillance Research, American Cancer Society, 250 Williams Street, Atlanta, GA 30303-1002 (e-mail: yiting-w@hotmail.com).
† Deceased.

(3, 4). Other researchers did not match or adjust for season of blood collection but used season-specific cutpoints to assign subjects to quartiles of 25(OH)D based on the season of the subject's sample collection (5). Yet others both matched on (and adjusted for) DOSC and used season-specific cutpoints to create categories (6).

To our knowledge, the impact of these different strategies on bias has not specifically been examined, and sometimes more than one of the above strategies was adopted in one study (7, 8). In particular, the rationale and impact of adjusting for DOSC deserves close examination. Prospectively, DOSC will generally not be associated with later disease risk and therefore would not usually be considered a confounding factor in nested case-control studies. However, it is known that if factors that influence nutrient levels are not taken into account, misclassification of the exposure may occur. If such a factor is not associated with disease, it contributes to random misclassification and attenuates any true biomarker-disease associations (9).

The purpose of this paper is to investigate the potential impact on bias associated with selected analytic strategies used in recently published studies to account for temporal variability in biomarker levels. We do not address impacts on statistical efficiency. We carried out simulations to explore whether adjustment for DOSC reduced bias, using the odds ratio as the measure of association between 25(OH)D (as an example of a biomarker that exhibits seasonal cyclic patterns) and a binary disease outcome. In addition, we compared various schemes of categorizing 25(OH)D that did, or did not, account for DOSC. Matching cases and controls on DOSC rather than just adjusting for DOSC is often used in epidemiologic studies. However, we focused on adjustment for DOSC in our simulations, without using matching, because matched selection of cases and controls (followed by appropriate adjustment for this matching) serves to make adjustment for DOSC merely more statistically efficient and is not expected to alter the magnitude of odds ratios (10). As noted below, this expectation was confirmed in selected simulations using matched selection of cases and controls.

We used 25(OH)D as a specific illustration of a biomarker with temporal variation, and we chose some of the parameters used in the simulations based on studies of 25(OH)D and its seasonal variation (7, 11–16). The magnitude of seasonal variation from a variety of geographic locations is generally close to or somewhat smaller than the value of 20 nmol/L (for highest vs. lowest 25(OH)D concentrations in a year) that we used in our main simulation. In addition, the results of our simulations could be generalized to other biomarkers with seasonal variation of varying magnitude similar to what we simulated for 25(OH)D. It is also possible that some of the patterns of bias observed in our simulations could be generalizable, at least qualitatively, to biomarkers with diurnal variation rather than seasonal variation.

MATERIALS AND METHODS

We used Monte Carlo simulations to evaluate bias in biomarker-disease odds-ratio estimates, using various analytic strategies. In these simulations, disease risk was determined solely by the true long-term average level, but short-term temporal changes affected the measured level. To make our simulations more concrete, we simulated exposure values that might be realistically observed if we were measuring serum 25(OH)D in units of nmol/L (1 ng/mL = 2.496 nmol/L). Data were generated with SAS version 9.1 software (SAS Institute, Inc., Cary, North Carolina), and macros are available upon request from the authors. This section of the paper describes 1) the assumed long-term average 25(OH)D levels and temporal changes, 2) the assumed effect of 25(OH)D on disease occurrence, 3) the analytic strategies, and 4) the simulation approach.

Model for individual 25(OH)D levels

Let μ_i be the true long-term average level of 25(OH)D for subject i that is causally responsible for the effect of 25(OH)D on the disease outcome. We refer to μ_i as the true level, and we assume it to be fixed for individual i. We also assume that true 25(OH)D concentrations in our simulated population follow a log-normal distribution, with a mean of $\boldsymbol{\mu}$ and a standard deviation of $\boldsymbol{\sigma}$.

With seasonal variation, the measured levels will fluctuate around μ_i. Let $m_i(t_i)$ be the measured 25(OH)D level for subject i from a single sample collected on calendar date t_i; we express $m_i(t_i)$ as the sum of 2 terms: the true level μ_i and the seasonal variation $f_i(t_i)$,

$$m_i(t_i) = \mu_i + f_i(t_i).$$

Here, $f_i(t_i)$ reflects the effect of temporal factors in terms of dates in a calendar year on the measured level; that is, this function of dates models the temporal variation. In our simulations, we assume that $f_i(t) = A_i \times \sin(t)$, where $\sin(t)$ is the trigonometric sine function to reflect sinusoidal changes over time, where the smallest unit of time in the present context is marked by dates in one calendar year, and A_i is the subject-specific amplitude of the variation. As noted in Table 1, the subject-specific A_i values were simulated by using a normal distribution. Therefore, in our simulations, each individual's 25(OH)D level at calendar date t_i is

$$m_i(t_i) = \mu_i + A_i \times \sin(t_i).$$

A sinusoidal function may be a plausible model for temporal variation not only for 25(OH)D but also for some other biomarkers. We assumed that the DOSC t_i for each subject was uniformly distributed throughout the year and was not associated with disease risk.

Model for 25(OH)D–disease association

Let π_i represent the disease risk for subject i over a certain time period of interest and $\pi_i|\mu_i$ the conditional probability of disease given that the subject's true 25(OH)D concentration is μ_i. We assume that the effect of 25(OH)D on disease risk is specified by the following model:

$$\text{logit}(\pi_i|\ \mu_i) = \alpha + \beta \times (\mu_i - \boldsymbol{\mu}),$$

where $\text{logit}(x)$ is the logistic function defined as $\text{logit}(x) = \ln[x/(1-x)]$ and α specifies the disease risk when the true

Table 1. Parameters Used in Simulations[a]

Parameter (Notation)	Specification and Coding
i	Index for each individual subject 1, 2, . . . , 100,000
True levels of 25(OH)D (μ_i)	Log-normal distribution with mean $\mu = 65$ and standard deviation $\sigma = 15$
Date of sample collection (t_i) for 25(OH)D	4 seasons and 12 months (uniform distribution on 0–2π):
	Spring (March–May): $0 < t_i \leq 0.25\pi$ (April 16–May 31) or $1.75\pi < t_i \leq 2\pi$ (March 1–April 15)
	Summer (June–August): $0.25\pi < t_i \leq 0.75\pi$
	Fall (September–November): $0.75\pi < t_i \leq 1.25\pi$
	Winter (December–February): $1.25\pi < t_i \leq 1.75\pi$
Amplitude of temporal variation in measured 25(OH)D (A_i)	Normal distribution with a mean of 20 and a standard deviation of 5
Temporal variation in measured 25(OH)D $f_i(t_i)$	$f_i(t_i) = A_i \times \sin(t_i)$
Parameters for the model of the association between 25(OH)D and disease (α and β)	$\alpha = -2.2$, $\beta = -0.22$
Odds ratio for every σ-unit increase in μ_i	Odds ratio $= \exp(\beta) = 0.8$

Abbreviation: 25(OH)D, serum 25-hydroxyvitamin D.

[a] Additional combinations of parameters are presented in the Web Appendix (posted on the *Journal*'s website: http://aje.oupjournals.org/).

25(OH)D concentration is at the population mean level $\boldsymbol{\mu}$. β is the increase in the log-odds ratio for each standard deviation increase in μ_i, and the risk odds ratio is $\exp(\beta)$. For example, when $\alpha = -2.2$ and $\beta = -0.22$, the risk of disease is 10% for a true level of $\mu_i = \boldsymbol{\mu}$, and the disease odds ratio for each $\boldsymbol{\sigma}$-unit increase in the true level (e.g., comparing a true level of $\mu_i + \boldsymbol{\sigma}$ with μ_i) is 0.8. We assumed that the odds ratio associated with each standard deviation increase in μ_i was constant across the range of simulated 25(OH)D concentrations. Note that disease risk does not depend on the measured level $m_i(t_i)$, given the true level μ_i.

We evaluate the situation that may often hold in practice: just one measured level $m_i(t_i)$ is available for each individual and is used (instead of μ_i) to estimate β.

Analytic strategies

To investigate how different analytic strategies impact bias, we applied several analytic strategies that have been used in published studies. In most analytic strategies, we categorized the measured 25(OH)D level into quartiles. The analytic strategies differed from one another in 2 main ways: first, how subjects were categorized into quartiles; second, by whether and how models were adjusted for DOSC. These strategies are described below. In an additional analysis, we treated the measured biomarker level as continuous (no categorization).

Categorizing exposure into quartiles. Subjects were categorized into quartiles of their measured 25(OH)D concentrations by using 3 different approaches. First, we used *overall* cutpoints *ignoring* DOSC and corresponding to the 25th, 50th, and 75th percentiles of measured concentrations in the entire population. Second, we used *season-specific* quartile cutpoints corresponding to the season-specific 25th, 50th, and 75th percentiles of measured levels in the population among participants who provided a sample in 1 of the 4 seasons (Table 1). Third, we used *month-specific* quartile cutpoints corresponding to the month-specific 25th, 50th, and 75th percentiles of measured levels in the population within each month.

Adjusting for temporal factors. We assessed the impact of considering DOSC by using 1 of 3 different analytic strategies and comparing each model with "truth": 1) no adjustment for DOSC, 2) adjustment for season of sample collection, and 3) adjustment for month of sample collection. The specific models that we fit are given subsequently in equation 1 (refer to the "Fitted model" section of the text below). Although DOSC is not a confounder in this situation because it is not associated with disease risk, some of the analytic strategies in the literature have included adjustment for DOSC, presumably to account for the effect of season on misclassification of the exposure, as previously noted.

Table 2 presents 8 models. Model 1 is a "gold standard" provided by fitting the specified model using the simulated true long-term average levels. Models 2–8 are based on simulated measured levels and are different combinations of approaches to defining quartile cutpoints (overall, season specific, or month specific) and of adjusting for DOSC (no adjustment, adjustment for season, and adjustment for month).

Treating the exposure as continuous. We also simulated treating the measured level as continuous and fitting the model

$$\text{logit}(\pi_i|\, m_i) = \text{logit}(\pi_i|\, m_i(t_i)) = \alpha^* + \beta^* \times m_i(t_i),$$

where $m_i(t_i)$ is the measured value of subject i. Similarly, we also compared models that adjusted for DOSC.

Model fitting and simulation methods

Fitted model. For each simulation, we fit the logistic model

$$\text{logit}(\pi_i|\, m_i) = a + \beta_2 \times \delta_{i2} + \beta_3 \times \delta_{i3} + \beta_4 \times \delta_{i4},$$

where δ_{ij} takes the value of 1 if subject i is in quartile j, and 0 otherwise. We estimated the βs from fitting the logistic

Table 2. Simulated Odds Ratios for the Association Between 25(OH)D and a Dichotomous Disease, Using Various Methods of Exposure Categorization and Adjustment for Date of Sample Collection

Model No.	25(OH)D Exposure Categorization	Adjustment for Date of Sample Collection	OR for Quartile 2	OR for Quartile 3	OR for Quartile 4 (High)
1	True levels ("gold standard")	None	0.806	0.703	0.569
2	Overall cutpoints, measured[a]	None	0.880	0.820	0.680
3	Overall cutpoints, measured[a]	Adjusted for season	0.811	0.700[b]	0.547[b]
4	Overall cutpoints, measured[a]	Adjusted for month	0.777[b]	0.644[b]	0.496[b]
5	Season-specific cutpoints, measured[c]	None	0.839	0.741	0.603
6	Season-specific cutpoints, measured[c]	Adjusted for season	0.839	0.741	0.603
7	Month-specific cutpoints, measured[c]	None	0.822	0.715	0.583
8	Month-specific cutpoints, measured[c]	Adjusted for month	0.822	0.715	0.582

Abbreviations: OR, odds ratio; 25(OH)D, serum 25-hydroxyvitamin D.

[a] Overall quartile cutpoints were obtained by pooling the measured 25(OH)D levels of all subjects, regardless of seasons or months of sample collection; refer to the Materials and Methods section of the text for more information.

[b] Indicates bias away from the null compared with odds-ratio estimates using true levels of 25(OH)D.

[c] Season-specific (or month-specific) cutpoints were obtained by assigning subjects to quartile categories within the same season (or month) of sample collection; refer to Materials and Methods for more information.

regression models with SAS (Proc Logistic) software. We adjusted for season or month in 2 ways:

1. Adding indicator variables for season to the model statement:

$$\text{logit}(\pi_i|\, m_i) = a + \beta_2 \times \delta_{i2} + \beta_3 \times \delta_{i3} + \beta_4 \times \delta_{i4} + \gamma_1 \times S_1 + \gamma_2 \times S_2 + \gamma_3 \times S_3. \quad (1)$$

Here, S_k (the indicator variable for season, where $_k = 1$, 2, or 3) is 1 if the sample was collected in season k, and 0 otherwise. Note that the subject index $_i$ is implicit for S_k but omitted for brevity.

2. Adding indicator variables for month to the model statement, analogous to equation 1 to adjust for month.

Simulation strategy. We considered different combinations of the simulation parameters, specified to reflect common scenarios. Table 1 shows the combination of parameters for which we present and discuss results in the text.

For each combination of simulation parameters, we generated a true 25(OH)D level for each of 100,000 subjects. We then determined the disease status for each subject based on his or her true 25(OH)D concentration. We compared a random number between 0 and 1 with the disease risk calculated from the model for the effect of the biomarker on disease risk (refer to the "Model for 25(OH)D–disease association" section of the text above). If the disease risk was greater than the random number, the disease status of the subject was considered diseased; otherwise, the subject was nondiseased. We also generated another random number between 0 and 2π to represent the DOSC. Without loss of generality, the interval between 0 and 2π was chosen to represent 1 year, and the 4 seasons and 12 months were represented by dividing this interval into 4 and 12 equal subintervals, respectively. The corresponding indicator variables for the seasons and months were determined by the subinterval in which the random date was categorized. We assumed that 25(OH)D levels peaked in midsummer (corresponding to $1/2\ \pi$) and were lowest in midwinter (corresponding to $3/2\ \pi$); therefore, the spring season was split into 2 parts (Table 1). The measured 25(OH)D level for each subject was then determined by using his or her true 25(OH)D level and DOSC, according to the model for 25(OH)D levels presented above (refer to the "Model for individual 25(OH)D levels" section of the text).

Logistic regression models specified above were fit and the odds-ratio estimates obtained. We repeated the process 1,000 times for each combination of assumed parameters (for each scenario). In this paper, we present the mean of the 1,000 resulting estimates to approximate the expected bias.

RESULTS

25(OH)D analyzed as a categorical variable

In Table 2, model 1 represents "truth," the odds ratio for each quartile of long-term average 25(OH)D level. When quartiles were created without accounting for DOSC (overall quartiles) and no adjustment was made for DOSC (model 2), odds ratios were biased toward the null compared with truth. In contrast, when overall quartiles were used but models were adjusted for either season (model 3) or month (model 4) of sample collection, odds ratios were biased away from the null (odds ratios footnoted in Table 2: associations were exaggerated compared with truth), although the amount of bias was not large.

When season-specific quartiles (model 5) or month-specific quartiles (model 7) were used, bias toward the null was reduced compared with use of overall quartiles (model 2), and there was no bias away from the null. When season-specific quartiles (model 6) or month-specific quartiles (model 8) were used, adjustment for season or month of blood collection had no effect on results.

Additional scenarios were considered, as indicated in the Web Appendix (Tables 1–15), and are summarized in Web

Table 3. Simulated Odds Ratios for the Association Between a Continuous Measure of 25(OH)D Level and a Dichotomous Disease, by Method of Adjusting for Date of Sample Collection

Adjustment for Date of Sample Collection	OR[a]
True levels ("gold standard")	0.796
Measured 25(OH)D, no adjustment for date of sample collection	0.895
Measured 25(OH)D, adjusted for season of sample collection	0.833
Measured 25(OH)D, adjusted for month of sample collection	0.810

Abbreviations: OR, odds ratio; 25(OH)D, serum 25-hydroxyvitamin D.
[a] Odds ratio associated with each 15 nmol/L increase in 25(OH)D level.

Appendix Table 2. (All 15 Web Appendix tables are posted on the *Journal*'s website (http://aje.oupjournals.org/).) In these scenarios, we examined the effect of changing values of the parameters for the amount of variation in 25(OH)D concentrations within the population, the degree of seasonal variation in 25(OH)D, the degree to which seasonal variation varied between individuals, and the size and direction of the association between 25(OH)D concentrations and the disease outcome. Some of these parameters would be expected to vary depending on the geographic location of the study population. For example, the amount of seasonal variation is likely to be larger at higher latitudes (larger A_i in Web Appendix Table 6 versus smaller A_i in Web Appendix Table 10), and the variation between individuals regarding the amount of seasonal variation is likely to be greater in populations that include people from across a large range of latitudes (a larger standard deviation in A_i in Web Appendix Table 9 versus a smaller standard deviation in A_i in Web Appendix Table 7). The overall patterns of biases remained similar in these additional scenarios, although the magnitude of bias varied. For example, a smaller standard deviation in the seasonal amplitude of change resulted in estimates from models using season-specific cutpoints that were closer to ''truth'' (Web Appendix Table 7 vs. Web Appendix Table 9). For another example, a very small seasonal amplitude while holding the other parameters constant also mitigated the magnitude of bias (Web Appendix Table 10 vs. Web Appendix Table 6).

25(OH)D analyzed as continuous

When 25(OH)D was analyzed as a continuous variable on its original scale and date of blood collection was ignored, odds ratios were biased toward the null, as expected (Table 3). However, adjustment for DOSC reduced this bias (Table 3). In no instance did adjustment for DOSC result in bias away from the null.

DISCUSSION

Several patterns emerged from our simulations, in which we applied 3 analytic strategies previously used in the literature to conditions that might plausibly be encountered in a study of blood 25(OH)D concentrations and a dichotomous disease outcome. First, as expected, ignoring DOSC resulted in bias toward the null. Second, using categories created without accounting for DOSC and then adjusting for season or month of blood collection resulted in bias away from the null in some simulations. Third, using categories that accounted for DOSC (e.g., season-specific or month-specific quartiles) did not result in bias away from the null but successfully reduced bias toward the null.

Contrary to our original expectations, we observed bias away from the null when 25(OH)D categories were created without accounting for DOSC and models were then adjusted for DOSC. Although the magnitude of bias away from the null that we found in our simulations was not large, this bias may be counterintuitive to many and deserves careful consideration.

The general concept that even nondifferential measurement error can cause bias away from the null has previously been documented (17–20). Flegal et al. (17) have shown that nondifferential measurement error in a continuous exposure measure can cause differential misclassification of exposure when the continuous exposure measure is categorized and cases have a different exposure distribution than noncases do (i.e., exposure is related to disease risk). In some Flegal et al. simulations, the resulting differential misclassification of exposure caused substantial bias away from the null when risk estimates were corrected by using methods appropriate for nondifferential misclassification (17). Our simulations differ from those of Flegal et al. in that we simulated measurement error caused by temporal variation (rather than purely random error), followed by modeling adjustment for the source of that temporal variation (DOSC).

Some empirical observations may help explain why bias away from the null occurred in some simulations when categorical variables for 25(OH)D were used. When quartiles were created without accounting for DOSC, the proportion of subjects within each quartile varied profoundly by season of sample collection. For example, of subjects with samples collected in winter, only about 0.5% were categorized into the highest quartile of 25(OH)D, while 57% were categorized into the lowest quartile of 25(OH)D. The 0.5% of subjects categorized into the highest quartile despite having a winter DOSC were those with extremely high true 25(OH)D concentrations; therefore, based on our disease risk model, they were at extremely low risk of disease. Conversely, of subjects with samples collected in summer, only 6.5% with extremely low 25(OH)D concentrations, and therefore at extremely high disease risk, were categorized into the lowest (referent) quartile of 25(OH)D, while about 61% of subjects were categorized into the highest quartile of 25(OH)D. Because risk estimates adjusted for season would reflect within-season risk estimates, models adjusted for season may be biased away from the null because risk estimates for those subjects with a winter or summer DOSC do not reflect a comparison of quartiles of true long-term 25(OH)D levels but rather a comparison of more extreme differences in 25(OH)D levels. We have referred to these odds-ratio estimates based on more extreme differences as being ''biased'' only because they differ from the odds ratios for quartiles of true long-term levels, which is how many might interpret odds ratios for quartiles. However, odds

ratios based on these more extreme differences could reasonably be considered biologically relevant measures, rather than biased.

Our simulations were not intended to answer the question of whether or not to match, because this question involves considerations other than bias, such as statistical efficiency (10). However, the implications of our simulation results should be the same for DOSC-matched analyses and for DOSC-adjusted analyses because both types of analysis yield the same expected value of the odds ratio (10). To illustrate, we repeated the analyses shown in Tables 2 and 3 by matching on DOSC (season/month) instead of adjusting for DOSC (simulated 1:1 matching of cases and controls on DOSC followed by conditional logistic regression analysis conditioning on the matched DOSC). As expected, these DOSC-matched analyses yielded results virtually identical to those from the DOSC-adjusted analyses (corresponding odds ratios always within 1%; data not shown, available on request). In DOSC-matched analyses, as in DOSC-adjusted analyses, using cutpoints that did not account for DOSC sometimes resulted in bias away from the null, whereas using DOSC-specific cutpoints did not.

One should consider how much we may generalize the observation of bias away from the null in our simulations. We note that meaningful bias away from the null did not always occur when adjustment for DOSC was used. For example, in Table 2, when models were adjusted for season of sample collection (model 3), odds ratios were very close to truth, presumably because biases away from and toward the null may have canceled each other out in this simulation. Only when adjustment for DOSC was more complete (for month of sample collection, model 4) was the net result a meaningful bias away from the null. More generally, we expect that the degree of bias away from the null, if any, will depend on the biomarker's distribution and temporal variation, the form and strength of the exposure-disease relation. Predicting the degree of bias away from the null (if any) in a particular analytic situation is likely to be complex. Nevertheless, our simulations illustrate that using categories created without accounting for DOSC and then adjusting for DOSC can result in bias away from the null under plausible conditions. Therefore, this strategy of dealing with measurement error from temporal variation can be less than ideal.

The alternative analytic strategy of accounting for DOSC when creating exposure categories (e.g., season- or month-specific categories) may be preferable. In our simulations, this strategy reduced bias toward the null and did not show any potential to introduce bias away from the null, as simple adjustment for DOSC did. Results of our simulations also suggest that if season- or month-specific categories are used, there is no additional advantage to further adjustment for DOSC since, as one would expect, it had no effect on the results.

It is worth noting that although using DOSC-specific exposure categories (e.g., season- or month-specific categories) was relatively effective in reducing bias in our simulations, it may be less effective under other conditions. In our simulation, every individual's 25(OH)D levels fluctuated from his or her true long-term level in a similar sinusoidal pattern over date, although we allowed the degree of fluctuation (the amplitude) to differ somewhat between individuals. Therefore, within short date intervals (e.g., a month), the approximate ranking of each study participant's 25(OH)D levels relative to those of other participants was preserved. In actuality, individuals' 25(OH)D levels may vary differently over the course of the year, depending on their seasonal patterns of outdoor activities and diet, for example. We found that use of DOSC-specific categories was less effective in reducing bias away from the null when the amount of temporal variation in biomarker levels varied widely between individuals (Web Appendix Tables 7–9).

Our model of the exposure-disease relation assumes that disease risk depends on each individual's long-term average level and that this relation does not depend on season of sample collection. Other exposure-disease models are possible but are beyond the scope of this paper. For example, disease risk might be strongly increased by transient or relatively short periods of vitamin D deficiency. In this situation, levels of 25(OH)D measured in the winter might be a stronger predictor of risk than levels measured during other seasons. Examining results stratified by season could be useful in exploring whether this type of exposure-disease relation appears to be present.

While we focused primarily on categorical measures of exposure, we also conducted simulations by using continuous exposure variables, which may yield the most informative estimates when a continuous model is a good fit for the true association between 25 (OH)D levels and disease risk.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology and Surveillance Research, American Cancer Society, Atlanta, Georgia (Yiting Wang, Eric J. Jacobs, Marjorie L. McCullough, Carmen Rodriguez, Michael J. Thun, Eugenia E. Calle); and Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia (W. Dana Flanders).

The authors thank Christina Newton at the American Cancer Society for her help with the SAS programs. They also thank Dr. Barry I. Graubard at the National Cancer Institute for sharing his valuable insights.

Conflict of interest: none declared.

REFERENCES

1. Tworoger SS, Hankinson SE. Use of biomarkers in epidemiologic studies: minimizing the influence of measurement error in the study design and analysis. *Cancer Causes Control.* 2006;17(7):889–899.
2. Webb AR. Who, what, where and when-influences on cutaneous vitamin D synthesis. *Prog Biophys Mol Biol.* 2006; 92(1):17–25.
3. Wu K, Feskanich D, Fuchs CS, et al. A nested case control study of plasma 25-hydroxyvitamin D concentrations and risk

of colorectal cancer. *J Natl Cancer Inst*. 2007;99(14): 1120–1129.
4. Feskanich D, Ma J, Fuchs CS, et al. Plasma vitamin D metabolites and risk of colorectal cancer in women. *Cancer Epidemiol Biomarkers Prev.* 2004;13(9):1502–1508.
5. Li H, Stampfer MJ, Hollis JB, et al. A prospective study of plasma vitamin D metabolites, vitamin D receptor polymorphisms, and prostate cancer [electronic article]. *PLoS Med*. 2007;4(3):e103.
6. Bertone-Johnson ER, Chen WY, Holick MF, et al. Plasma 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2005; 14(8):1991–1997.
7. Stolzenberg-Solomon RZ, Vieth R, Azad A, et al. A prospective nested case-control study of vitamin D status and pancreatic cancer risk in male smokers. *Cancer Res.* 2006; 66(20):10213–10219.
8. Freedman DM, Chang SC, Falk RT, et al. Serum levels of vitamin D metabolites and breast cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial. *Cancer Epidemiol Biomarkers Prev.* 2008;17(4):889–894.
9. Hunter D. Biochemical indicators of dietary intake. In: Willett WC, ed. *Nutritional Epidemiology.* 2nd ed. New York, NY: Oxford University Press; 1998:174–243.
10. Rothman KJ, Greenland S, eds. *Modern Epidemiology.* 2nd ed. Philadelphia, PA: Lippincott-Raven; 1998:147–300.
11. Bischoff-Ferrari HA. The 25-hydroxyvitamin D threshold for better health. *J Steroid Biochem Mol Biol.* 2007;103(3–5): 614–619.
12. Harris SS, Dawson-Hughes B. Seasonal changes in plasma 25-hydroxyvitamin D concentrations of young American black and white women. *Am J Clin Nutr.* 1998;67(6): 1232–1236.
13. Munger KL, Levin LI, Hollis BW, et al. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. *JAMA*. 2006; 296(23):2832–2838.
14. Oliveri MB, Ladizesky M, Mautalen CA, et al. Seasonal variations of 25 hydroxyvitamin D and parathyroid hormone in Ushuaia (Argentina), the southernmost city of the world. *Bone Miner.* 1993;20(1):99–108.
15. Stamp TC, Round JM. Seasonal changes in human plasma levels of 25-hydroxyvitamin D. *Nature*. 1974;247(442):563–565.
16. Woitge HW, Knothe A, Witte K, et al. Circaannual rhythms and interactions of vitamin D metabolites, parathyroid hormone, and biochemical markers of skeletal homeostasis: a prospective study. *J Bone Miner Res.* 2000;15(12):2443–2450.
17. Flegal KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. *Am J Epidemiol*. 1991;134(10):1233–1244.
18. Birkett NJ. Effect of nondifferential misclassification on estimates of odds ratios with multiple levels of exposure. *Am J Epidemiol*. 1992;136(3):356–362.
19. Dosemeci M, Wacholder S, Lubin JH. Does nondifferential misclassification of exposure always bias a true effect toward the null value? *Am J Epidemiol*. 1990;132(4):746–748.
20. Wacholder S. When measurement errors correlate with truth: surprising effects of nondifferential misclassification. *Epidemiology.* 1995;6(2):157–161.