Research Article

Statistics in Medicine

Received 5 April 2015, Accepted 2 December 2015 Published online 5 January 2016 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6856

Calibration and seasonal adjustment for matched case–control studies of vitamin D and cancer

Mitchell H. Gail,[a*†] Jincao Wu,[b] Molin Wang,[c,d] Shiaw-Shyuan Yaun,[e] Nancy R. Cook,[d,f] A. Heather Eliassen,[d,g] Marjorie L. McCullough,[h] Kai Yu,[a] Anne Zeleniuch-Jacquotte,[i] Stephanie A. Smith-Warner,[d,e] Regina G. Ziegler[a] and Raymond J. Carroll[j,k]

Vitamin D measurements are influenced by seasonal variation and specific assay used. Motivated by multicenter studies of associations of vitamin D with cancer, we formulated an analytic framework for matched case–control data that accounts for seasonal variation and calibrates to a reference assay. Calibration data were obtained from controls sampled within decile strata of the uncalibrated vitamin D values. Seasonal sine–cosine series were fit to control data. Practical findings included the following: (1) failure to adjust for season and calibrate increased variance, bias, and mean square error and (2) analysis of continuous vitamin D requires a variance adjustment for variation in the calibration estimate. An advantage of the continuous linear risk model is that results are independent of the reference date for seasonal adjustment. (3) For categorical risk models, procedures based on categorizing the seasonally adjusted and calibrated vitamin D have near nominal operating characteristics; estimates of log odds ratios are not robust to choice of seasonal reference date, however. Thus, public health recommendations based on categories of vitamin D should also define the time of year to which they refer. This work supports the use of simple methods for calibration and seasonal adjustment and is informing analytic approaches for the multicenter Vitamin D Pooling Project for Breast and Colorectal Cancer. Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA.

Keywords: calibration; seasonal adjustment; measurement error; matched case–control study; molecular epidemiology; biomarkers

1. Introduction

Measurements of 25-hydroxyvitamin D, which we call vitamin D, are influenced by seasonal variation and assay calibration. Vitamin D blood concentrations increase in response to sun exposure, inducing seasonal changes. These factors need to be taken into account in the analysis of matched case–control

[a]*Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, U.S.A.*
[b]*Division of Biostatistics, Center for Devices and Radiological Health, Food and Drug Administration, Silver Spring, MD, U.S.A.*
[c]*Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, U.S.A.*
[d]*Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, U.S.A.*
[e]*Department of Nutrition, Harvard T. H. Chan School of Public Health, Boston, MA, U.S.A.*
[f]*Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, U.S.A.*
[g]*Channing Division of Network Medicine, Department of Medicine, Brigham and Womens Hospital and Harvard Medical School, Boston, MA, U.S.A.*
[h]*Epidemiology Research Program, American Cancer Society, Atlanta, GA, U.S.A.*
[i]*Department of Population Health and Perlmutter Cancer Center, New York University School of Medicine, New York, NY, U.S.A.*
[j]*Department of Statistics, Texas A&M University, College Station, TX, U.S.A.*
[k]*Department of Mathematics and Statistics, University of Technology, Sydney, Australia*
**Correspondence to: Mitchell H. Gail, Biostatistics Branch, National Cancer Institute, 9609 Medical Center Drive, Rm. 7E-138, Rockville, MD, 20850-9780, U.S.A.*
[†]*E-mail: gailm@mail.nih.gov*

studies to associate vitamin D with disease risk. Although previous case–control analyses have used sine–cosine series to account for seasonal variation [1, 2], as we do, we have not found studies that allow for both calibration and seasonal adjustment. In this paper, we present a framework for such analyses. This framework, and parameters based on real studies, underlie our simulations to evaluate the performance of various procedures to estimate log odds ratios.

This work was motivated by collaboration on the Vitamin D Pooling Project of Breast and Colorectal Cancer (hereafter, Vitamin D Pooling Project), which includes 21 cohort studies in North America, Europe, and Asia. Because vitamin D is stable in stored frozen blood samples, such samples can be used to study associations with cancers that develop years later. Vitamin D was measured in previously stored blood from incident breast or colon cancer cases in these cohorts and from their matched controls. In some studies, the controls were tightly matched to cases on date of blood draw, but not in all studies. Moreover, different assays were used in various studies. Hence, calibration against a reference laboratory was required to put measurements from various studies on a common scale. In each study, reference laboratory measurements were obtained from 29 control bloods, selected by stratified random sampling within strata defined by deciles of the uncalibrated control vitamin D measurements. To control for effects of seasonal variation in within-study comparisons of cases and controls, study-specific seasonal adjustment was required. The intent of using calibration and study-specific seasonal adjustment is to transform the original data so that every measurement, regardless of when it was drawn and regardless of study assay, could be thought of as having been measured on the same reference date by the reference laboratory.

This work has several novel features. It provides a statistical framework to accommodate both calibration and seasonal adjustment. Using this framework, we assess procedures for inference on log odds ratios, both for continuous and categorical vitamin D risk models, in realistic simulations based on data from the Vitamin D Pooling Project. Finally, we identify practical recommendations for analysis and interpretation to inform the work of the Vitamin D Pooling Project. In particular, our analyses show the following: (1) failure to adjust for calibration and seasonal trend can inflate variance, bias, and mean square error of estimated vitamin D effects; (2) simple analytic methods can be recommended for continuous and categorical risk models (and perform even better than a theoretically appealing normal model for categorical risk); (3) with the stratified sampling design for calibration samples, there is surprisingly little increase in the variance of log odds ratio estimates from calibration and seasonal adjustment; and (4) unlike continuous vitamin D risk models, categorical models are not robust to the choice of reference date for seasonal adjustment. This finding implies that public heath recommendations for desirable vitamin D levels may need to be season specific.

Section 2 describes models and statistical methods. Section 3 describes simulation studies in which the logit of disease risk is linear or categorical in vitamin D. Section 4 presents analyses of data from three nested case–control studies, and Section 5 has concluding remarks. Technical results are given in an Appendix, and additional methods, simulations, and examples are given in the supporting information.

2. Methods

2.1. Motivating data

We base our simulations (Section 3) on two typical data sets, one a nested case–control study of colorectal cancer from the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study [3], hereafter ATBC, and the second, a nested case–control study of breast cancer from the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial or PLCO [4]. We use only data from controls from these studies as the basis of our simulations, and we specify the exposure effects. Figure 1 depicts vitamin D measurements in ATBC controls as a function of the week the blood sample was drawn ranging from week 1 (January 1–7) to week 52. The solid line represents a sine/cosine series fit to these data, described fully in Section 2.3. No samples were collected between weeks 24 and 32. The maximum value of the fitted curve is at week 35 (the third week in September) and the fitted values at the beginning and end of the year agree, as would make sense because they represent the same time of year. Another plot of this type is given for the PLCO data as Figure S1. In addition to the vitamin D measurements that are available from the study-specific laboratory on all subjects, we have calibration measurements from a reference laboratory on a sample of 29 controls, selected as described in Section 2.4.

We also estimated disease associations with vitamin D from cases and controls in three studies, ATBC (Table V), and two studies of breast cancer, the New York University Women's Health Study (NYUWHS) (Table SX) and the Cancer Prevention Study II Nutrition Cohort (CPSII) (Table SXI).

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

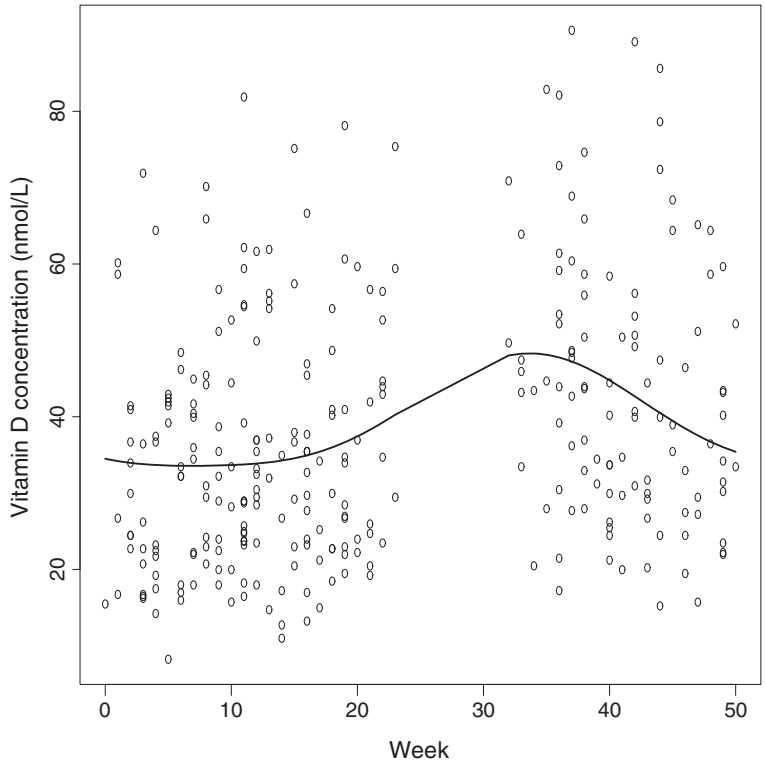

Figure 1. Seasonal trends in vitamin D measurements in controls from the Alpha-Tocopherol, Beta-Carotene study.

Descriptive data for these four cohorts are in Table SI.

2.2. *Statistical framework for calibration and seasonal trend models*

The following notation is for a single study. Let $w(t) = \varsigma(t) + w$ be the true value of the study-specific (or 'local') assay for an individual measured at week t, where $\varsigma(t)$ is the seasonal trend in the population and $w \sim \text{Normal}(0, \tau^2)$ represents the deviation of the individual's value from the trend. The observed local assay value is $W(t) = w(t) + u_W$, where $u_W \sim \text{Normal}(0, \sigma_{WW})$ represents measurement error. We assumed a linear calibration curve because the regression of reference vitamin D measurements against local laboratory measurements was usually linear in Vitamin D Pooling Project studies, and in those few instances where a quadratic component was detected, differences in predictions between the linear and quadratic models were small. Under linear calibration, we denote the true value for the reference laboratory as $x(t) = a + bw(t)$, where a and b represent the true calibration intercept and slope. The observed reference laboratory measurement would be $X(t) = x(t) + u_X$, where $u_X \sim \text{Normal}(0, \sigma_{XX})$ is measurement error. Assuming that $(w, u_W, u_X)^{\mathrm{T}}$ are trivariate normal with $\text{cov}(w, u_X) = \text{cov}(w, u_W) = 0$ and $\text{cov}(u_X, u_W) = \sigma_{XW}$, it follows that the observable variables $\{W(t), X(t)\}^{\mathrm{T}}$ are bivariate normal with means $\{\varsigma(t), a + b\varsigma(t)\}^{\mathrm{T}}$ and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \tau^2 + \sigma_{WW} & b\tau^2 + \sigma_{XW} \\ b\tau^2 + \sigma_{XW} & b^2\tau^2 + \sigma_{XX} \end{pmatrix}.$$

Hence, the regression of $X(t)$ on $W(t)$ is

$$\begin{aligned} E\{X(t)|W(t)\} &= a + b\varsigma(t) + (b\tau^2 + \sigma_{XW})(\tau^2 + \sigma_{WW})^{-1}\{W(t) - \varsigma(t)\} \\ &= a + \varsigma(t)\left\{b - (b\tau^2 + \sigma_{XW})(\tau^2 + \sigma_{WW})^{-1}\right\} \\ &\quad + (b\tau^2 + \sigma_{XW})(\tau^2 + \sigma_{WW})^{-1}W(t). \end{aligned} \tag{1}$$

If we assume that $\sigma_{XW} = 0$, as is reasonable because measurements in the local and reference laboratories on the same sample are usually performed in different locations and/or times, the regression reduces to

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

$$E\{X(t)|W(t)\} = E\{x(t)|W(t)\} = a + b\varsigma(t) + b\kappa_W\{W(t) - \varsigma(t)\}, \quad (2)$$

where $\kappa_W = \tau^2(\tau^2 + \sigma_{WW})^{-1}$ is an attenuation factor (or intraclass correlation) from measurement error in $W(t)$. Local assay values $W(t)$ are available on all samples, whereas reference laboratory values $X(t)$ are available only on a sample of controls, as described in Section 2.4. We assume σ_{XX} and σ_{WW} are known from reliability studies in which the same sample is repeatedly analyzed on different days in the reference and local laboratories. Data from the studies in the Vitamin D Pooling Project indicate that $(\sigma_{XX})^{1/2}$ and $(\sigma_{WW})^{1/2}$ each ranged from 0.5 to 9 nmol/L, with most values in the range of 2–4. To estimate τ^2 and hence κ_W, we subtracted σ_{WW} from the empirical variance of $\{W(t) - \widehat{\varsigma}(t)\}$ in controls, where $\widehat{\varsigma}(t)$ is the estimated periodic trend in $W(t)$. Because τ^2 was much larger than σ_{WW}, results were little changed in sensitivity analyses as σ_{XX} and σ_{WW} ranged from 4 to 36. In Section 2.3, we describe how we fit the trend.

2.3. *Fitting the trend*

We used data on $W(t)$ from controls to estimate the seasonal trend $\varsigma(t)$ by fitting the model

$$W(t) = \varsigma(t) + \varepsilon, \quad (3)$$

$$\varsigma(t) = \gamma_0 + \gamma_1 \sin(2\pi t/52) + \gamma_2 \cos(2\pi t/52) + \gamma_3 \sin(4\pi t/52) + \gamma_4 \cos(4\pi t/52), \quad (4)$$

where $\varepsilon \sim \text{Normal}(0, \sigma^2)$ is independent of $\varsigma(t)$ and t is in weeks. Such periodic series have been studied previously (e.g., [5]) and shown to fit vitamin D data as an outcome variable with t in days [6] and as an exposure in case–control analyses with t in months [1, 2]. Those vitamin D analyses used only the first three terms in Equation (4), but we found that five terms were needed to fit the data well in each of the 21 Vitamin D Pooling Project studies. Figure 1 shows the fit to data from the controls in the ATBC colorectal case–control study. This approach has the advantages that only 5 degrees of freedom are used to fit the trend, which has little impact on the variance of risk estimates (Supporting Information Appendices A.5 and A.6), and the trend gives similar values on December 31 as on January 1, as is desirable. When we tried other more flexible fitting methods, such as splines with many knots or LOESS regression, the trends were not necessarily equal at the beginning and end of the year, and the variance of estimates of vitamin D effects were inflated, as previously noted [7]. Controls were used for fitting seasonal trends because, for low-incidence diseases like breast or colorectal cancer, controls are representative of the general population (see the rare disease assumption in the Supporting Information Appendix A.1). Matching controls to cases on cancer risk factors such as age changes the age distribution in controls but would not alter the seasonal vitamin D pattern in controls appreciably unless this was strongly associated with age.

2.4. *Estimating the calibration parameters* (*a and b*)

The Vitamin D Pooling Project included calibration samples from controls within each study. For each study, the control values of $W(t)$ were grouped into deciles, and from each decile, three controls were selected at random. (In fact, 29 controls were usually selected this way, rather than 30, but for the simulations later, we selected three from each decile.) Blood samples from the selected controls were then sent to the reference laboratory for measurement of $X(t)$. Thus, the $\{X(t), W(t)\}$ pairs are a stratified random sample with strata defined by deciles of $W(t)$.

We considered three estimates of the calibration parameters a and b.

- Estimates $\widehat{a}_1$ and $\widehat{b}_1$ are obtained from simple linear regression of $X(t)$ on $W(t)$ in the calibration data. These quantities are not consistent for a and b, in view of Equation (2).
- From Equation (2), a somewhat more refined estimate $(\widehat{a}_2, \widehat{b}_2)^{\mathrm{T}}$ is obtained by regressing $X(t)$ on $\widehat{\kappa}_W\{W(t) - \widehat{\varsigma}(t)\}$.
- The most justifiable estimate, in view of Equation (2), is $(\widehat{a}_3, \widehat{b}_3)^{\mathrm{T}}$, which is obtained by regressing $X(t)$ on $\widehat{\varsigma}(t)(1 - \widehat{\kappa}_W) + \widehat{\kappa}_W W(t)$. For $\widehat{\kappa}_W$ equal 1, this procedure reduces to the regression of $X(t)$ on $W(t)$.

We evaluated each of these approaches in simulations.

2.5. *Risk models and estimators*

2.5.1. Continuous risk models. For modeling continuous vitamin D, we assumed that the logit of the disease risk is

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

$$\text{logit}\{\text{pr}(Y=1)\} = \alpha_0 + \alpha_1 x_0, \tag{5}$$

where $x_0 \equiv x(t_0) = a + bw(t_0)$ is the value that the true reference laboratory would have produced if it had been measured on the reference date t_0 instead of the date t. Each matched case–control pair makes the following contribution to the conditional likelihood:

$$\exp\{\alpha_1(x_{0,\text{case}} - x_{0,\text{control}})\}/[\exp\{\alpha_1(x_{0,\text{case}} - x_{0,\text{control}})\} + 1]. \tag{6}$$

However, we only get to observe $W(t)$, not x_0. One approach is to replace x_0 by its conditional expectation given $W(t)$ in Equation (2), which we estimate by

$$\widehat{x}_0 = \widehat{a} + \widehat{b}\widehat{\varsigma}(t_0) + \widehat{b}\widehat{\kappa}_W\{W(t) - \widehat{\varsigma}(t)\} = \widehat{a} + \widehat{b}\{\widehat{\varsigma}(t_0) - \widehat{\kappa}_W\widehat{\varsigma}(t)\} + \widehat{b}\widehat{\kappa}_W W(t). \tag{7}$$

The estimates $\widehat{\varsigma}(\cdot)$, $\widehat{a}$, and $\widehat{b}$ are described in Sections 2.3 and 2.4. Then (7) leads to the estimated difference

$$\widehat{x}_{0,\text{case}} - \widehat{x}_{0,\text{control}} = \widehat{b}\widehat{\kappa}_W\{W(t_{\text{case}}) - W(t_{\text{cont}}) - \widehat{\varsigma}(t_{\text{case}}) + \widehat{\varsigma}(t_{\text{cont}})\}. \tag{8}$$

Under the rare disease assumption, and using Equation (8), if one substitutes $\{W(t_{\text{case}}) - W(t_{\text{cont}}) - \widehat{\varsigma}(t_{\text{case}}) + \widehat{\varsigma}(t_{\text{cont}})\}$ for $(x_{0,\text{case}} - x_{0,\text{control}})$ in Equation (6) and maximizes the conditional likelihood with respect to α_1, one obtains the estimate $\widehat{\alpha}_1^*$ of $\alpha_1^* = \alpha_1 b\kappa_W$. Thus, we estimate α_1 by $\widehat{\alpha}_1 = \widehat{\alpha}_1^*/(\widehat{b}\widehat{\kappa}_W)$.

The asymptotic behavior of $\widehat{\alpha}_1$ is derived in the Supporting Information Appendix A.6. A simple estimate of its variance is obtained as follows. First, estimate the variance of $\widehat{\alpha}_1^*$ by the inverse of the Hessian of the conditional log-likelihood, and call this $\widehat{V}\text{ar}(\widehat{\alpha}_1^*)$. Then, by the delta method, $\widehat{\alpha}_1$ has approximate variance estimate $\widehat{V}\text{ar}(\widehat{\alpha}_1) = (\widehat{b}\widehat{\kappa}_W)^{-2}\widehat{V}\text{ar}(\widehat{\alpha}_1^*) + (\widehat{\alpha}_1^*/\widehat{b}^2\widehat{\kappa}_W)^2\widehat{V}\text{ar}(\widehat{b})$ where $\widehat{V}\text{ar}(\widehat{b})$ is estimated from the regressions in Section 2.4. In this calculation, we use the fact that $\widehat{\alpha}_1^*$ and $\widehat{b}$ are independent, and we assume that the variation from estimates of trend and $\widehat{\kappa}_W$ is numerically negligible (Supporting Information Appendix A.6). Some intuition is obtained by considering the ATBC example. Even if the case and control had $t_{\text{case}} = t_{\text{cont}}$, $V\text{ar}[\{W(t_{\text{case}}) - W(t_{\text{cont}})\}] \geqslant 2\tau^2 = 2 \times 240.9 = 481.8$, which greatly exceeds the variance of the difference in trend estimates. For example, $V\text{ar}\{\widehat{\varsigma}(39/52) - \widehat{\varsigma}(1/52)\} = 10.9$, which is much smaller than 481.8. Likewise, the standard error of $\widehat{\kappa}_W$ was 0.0066 in simulations, implying small variation about $\kappa_W = 0.9377$ with $\sigma_{WW} = 16$.

Importantly, none of these calculations depend on the calibration intercept a. The seasonal adjustment in Equation (8) vanishes if the case and control are matched on the date of blood draw so that $t_{\text{case}} = t_{\text{cont}}$. Also, the seasonal adjustment in (8) does not depend on the reference date, t_0. Ignoring calibration and measurement error in $W(t)$ is equivalent to using $\widehat{\alpha}_1^*$ instead of $\widehat{\alpha}_1$.

2.5.2. Categorized risk models. For categorical vitamin D, we assume

$$\text{logit}\{P(Y=1)\} = \mu + \sum_{i=1}^{5} I(c_{i-1} \leqslant x_0 < c_i)\beta_i, \tag{9}$$

where $c_0 = 0$ and $\beta_1 = 0$. For 1:1 nested case–control matching, the corresponding conditional likelihood contribution from a case–control pair is

$$\exp\left\{\boldsymbol{\beta}^T C_1(x_0)\right\} / \left[\exp\left\{\boldsymbol{\beta}^T C_1(x_0)\right\} + \exp\left\{\boldsymbol{\beta}^T C_0(x_0)\right\}\right], \tag{10}$$

where $C_\ell(x)$ is a 5×1 vector of indicators in (9) for cases ($\ell = 1$) and controls ($\ell = 0$) and $\boldsymbol{\beta}$ is the corresponding vector of log odds ratios, β_i. A linear trend risk model sets $\beta_i = (i-1)\beta$, where β is the log odds increase in risk per exposure category increase. The contribution to the likelihood for the linear model is Equation (10) with $(i-1)\beta$ replacing $\boldsymbol{\beta}^T C_\ell(x_0)$. The cut-points c_i could be externally defined levels, such as those given by the Institute of Medicine for vitamin D [8]. They could also be study-wide quantiles of vitamin D estimated from controls from all participating studies. In the simulations in Section 3.3, they were quintiles of x_0 values among controls from the particular study on which the simulations were based. In particular, we set $x_0 = a + b\{W(t) - \widehat{\varsigma}(t) + \widehat{\varsigma}(t_0)\}$ where $W(t)$ and $\widehat{\varsigma}(\cdot)$ were from controls from that study.

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

If x_0 were known, we could estimate $\boldsymbol{\beta}$ or β by maximizing the product of conditional likelihoods (10) with respect to $\boldsymbol{\beta}$ or β and obtain model-based estimates of their covariance from the Hessian of the log-likelihood. With $\hat{x}_0$ defined in (7), we proceeded by substituting $\hat{x}_0$ for x_0 in Equation (10) and by estimating covariances from the Hessian as if x_0 were known. We investigated the operating characteristics of this approach in simulations.

Although the preceding estimation procedures are easy to describe and implement, they do not take full advantage of the assumed normal distributions, nor do they acknowledge the uncertainty in assigning risk category based on $\hat{x}_0$ instead of x_0. Given $a, b,\ \varsigma(\cdot)$, and $W(t), x_0 \sim \text{Normal}[a+b\varsigma(t_0)+b\kappa_W\{W(t)-\varsigma(t)\},\ \eta^2 \equiv b^2\sigma_{WW}\kappa_W]$. Hence,

$$\text{pr}\left\{x_0 \text{ in category } i|W(t)\right\} \equiv E\left\{I(c_{i-1} \leqslant x_0 < c_i)|a, b, \varsigma(\cdot), W(t)\right\} = \Phi(z_i) - \Phi(z_{i-1}), \tag{11}$$

where $z_i = [c_i - a - b\varsigma(t_0) - b\kappa_W\{W(t) - \varsigma(t)\}]/\eta$ and $\Phi(\cdot)$ is the standard normal distribution function. For a rare disease, the conditional likelihood corresponding to a matched case–control set is shown in Appendix A.1 to be

$$\frac{\sum_{i=1}^{5} \exp(\beta_i)\text{pr}\{x_{0,\text{case}} \text{ in category } i|W(t_{\text{case}})\}}{\sum_{i=1}^{5} \exp(\beta_i)\text{pr}\{x_{0,\text{case}} \text{ in category } i|W(t_{\text{case}})\} + \sum_{i=1}^{5} \exp(\beta_i)\text{pr}\left\{x_{0,\text{control}} \text{ in category } i|W(t_{\text{cont}})\right\}}. \tag{12}$$

The product of such terms over case–control pairs can be maximized with respect to $\boldsymbol{\beta}$ in the general model or β (the per category log odds ratio) in the linear trend model to obtain estimates and their model-based estimated covariance. We substituted estimates $\hat{\varsigma}(\cdot)$, $\hat{a}$, and $\hat{b}$ to compute $\text{pr}\{x_0 \text{ in } i|W(t)\}$ in Equation (11). A more precise analysis might maximize the average likelihood over the distribution of $\hat{\varsigma}(\cdot)$, $\hat{a}$, and $\hat{b}$. A numerical approach would be to draw repeated samples of $\hat{\varsigma}(\cdot)$, $\hat{a}$, and $\hat{b}$ to compute an average likelihood and maximize this quantity, but we have not pursued it because unreported simulations indicated that substitution of $\hat{\varsigma}(\cdot)$, $\hat{a}$, and $\hat{b}$ yields similar operating characteristics as using the exact quantities $\varsigma(\cdot)$, a, and b.

3. Simulations

3.1. Simulation setup

In this section, we describe methods to simulate realistic case–control samples based on controls from the colorectal cancer case–control study in ATBC. We also used these methods for simulations based on the breast cancer case–cohort study in PLCO. We describe the analytic procedures for continuous and categorical risk models.

Data from the controls were used to estimate the distribution of times of blood draw, the trend $\varsigma(t)$, and τ^2. In particular we estimated $\tau^2 = 240.89$ assuming $\sigma_{WW} = 16$. The estimated coefficients in the seasonal trend Equation (4) were 39.425, −5.837, −4.485, 1.440, and −0.444.

We generated 10,000 case–control studies based on the original ATBC data. For each simulated study, we generated a large ($N = 50{,}000$) source population as follows. First, draw a date t of blood by sampling with replacement from the dates in the original controls. Then, generate $w(t) = \varsigma(t)+w$ by sampling from $w \sim \text{Normal}(0, \tau^2)$. Set $x_0 = a+bw(t_0) = a+b\{\varsigma(t_0)+w\}$ and generate the disease status indicator Y from Equation (5) for continuous vitamin D and from Equation (9) for categorical vitamin D. Generate the observed local vitamin D measurement from $W(t) = w(t) + u_W$ by drawing from $u_W \sim \text{Normal}(0, \sigma_{WW})$, where σ_{WW} is assumed known, and independently generate reference laboratory measurements $X(t) = x(t) + u_X = a + bw(t) + u_X$ by drawing from $u_X \sim \text{Normal}(0, \sigma_{XX})$ with σ_{XX} known. Implicitly, we are assuming $\sigma_{XW} = 0$. At this stage, we have N triples $\{Y, W(t), X(t)\}^T$.

To obtain the case–control data, we randomly selected n triples from among the triples with $Y = 1$ (cases) and n triples from among those with $Y = 0$ (controls). However, we observed $X(t)$ only in the calibration sample, which is obtained by arranging the $W(t)$ values of controls into decile groups and selecting three controls at random from among the controls in each decile group. For these 30 controls, which we call the calibration sample, we observed $\{Y, W(t), X(t)\}^T$, but in all other cases and controls, we only observed $\{Y, W(t)\}^T$. These triplets and doublets constituted the case–control data available for analysis. The effects of matching variables on risk cancel from the conditional logistic likelihood. We

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

assumed that the matching variables have no effect on risk in these simulations, because there are no matching effects in Equations (5) and (9). Hence, we randomly matched cases with controls to obtain n pairs, regardless of season.

Analysis of each such simulated data set for continuous vitamin D proceeded as follows. First, obtain $\widehat{\varsigma}(t)$ by fitting Equation (3) to the n control values of $W(t)$. Next estimate τ^2 by subtracting σ_{WW} from the variance of $W(t)$ - $\widehat{\varsigma}(t)$ in controls. Use $\widehat{\tau}^2$ and the assumed known value σ_{WW} to calculate $\widehat{\kappa}_W$. Let $\widehat{\alpha}_1^*$ be the estimate obtained from conditional logistic regression (CLR) using the residuals $\{W(t) - \widehat{\varsigma}(t)\}$. We described three estimates of $(a,b)^T$ in Section 2.4 and used these results to produce three estimates of α_1, namely, $\widehat{\alpha}_1 = \widehat{\alpha}_1^*/\widehat{b}\widehat{\kappa}_W$, with $\widehat{b}$ being $\widehat{b}_1$, $\widehat{b}_2$, or $\widehat{b}_3$. We also present estimates from CLR with raw values $W(t)$ and with only seasonally adjusted values, $\{W(t) - \widehat{\varsigma}(t)\}$. Finally, we present an estimate based on calibration only that is obtained by dividing the estimate from the raw values, $W(t)$, by $\widehat{b}_1\widehat{\kappa}_W$. All these estimates were compared with estimates obtained by inserting the true values x_0 from the case and matched control into the CLR. None of these estimates depends on a standard reference date, t_0.

For categorical vitamin D, values of Y were generated from Equation (9) based on quintiles of x_0 from the study on which the simulation was based, as described in Section 2.5.2 . A benchmark analysis used the perfect categories in Equation (10) based on the true x_0, which is unknown in practice. We also categorized the raw values $W(t)$ without seasonal adjustment or calibration for use in Equation (10), as well as the seasonally adjusted but not calibrated values, $W(t) - \widehat{\varsigma}(t) + \widehat{\varsigma}(t_0)$ and the calibrated but not seasonally adjusted values $\widehat{a}_1 + \widehat{b}_1 W(t)$. Two more principled estimates $\widehat{x}_0$ were used for categorization in Equation (10). We call them $\widehat{x}_{01} = \widehat{a}_1 + \widehat{b}_1\widehat{\varsigma}(t_0) + \widehat{b}_1\widehat{\kappa}_W\{W(t) - \widehat{\varsigma}(t)\}$ and $\widehat{x}_{03} = \widehat{a}_3 + \widehat{b}_3\widehat{\varsigma}(t_0) + \widehat{b}_3\widehat{\kappa}_W\{W(t) - \widehat{\varsigma}(t)\}$. A final estimate of log odds ratios, called the 'normal model' estimate, was based on Equations (11) and (12) with $\widehat{a}_3 + \widehat{b}_3\widehat{\varsigma}(t_0) + \widehat{b}_3\widehat{\kappa}_W\{W(t) - \widehat{\varsigma}(t)\}$ and $\widehat{\eta} = \widehat{b}_3(\sigma_{WW}\widehat{\kappa}_W)^{1/2}$ substituted for corresponding parameters to compute z_i.

3.2. Analysis of continuous vitamin D

In this section, we compare procedures with calibration and/or seasonal adjustment for analyzing continuous vitamin D. Table I contains comparisons for the various estimators of slope under the null hypothesis $\alpha_1 = 0$. The logistic intercept was $\mu = -4$. The parameters were chosen to provide a severe test of the various procedures. In particular, there were only $n = 100$ case–control pairs; larger numbers of n make it easier to estimate trends and reduce small sample bias in estimators. The values $a = 5$ and $b = 1.4$ represent more severe miscalibration of the local assay than found in most Vitamin D Pooling Project studies. The estimated measurement error variances $\sigma_{XX} = \sigma_{WW} = 16$ are representative of values found in replication

Table I. Simulation performance of procedures for conditional logistic regression analysis of continuous vitamin D based on data from the Alpha-Tocopherol, Beta-Carotene study at the null hypothesis $\alpha_1 = 0$, with $n = 100$ case–control pairs.

	Bias of $\widehat{\alpha}_1$ ($\times 10^5$)	CI%	SE ($\widehat{\alpha}_1$) ($\times 10^2$)	SE ratio	Var ratio	MSE ratio
Exposure x measured without error	7.24	94.51	0.687	1.032	1.000	1.000
Calibration/season: estimate $(\widehat{a}_1, \widehat{b}_1)$ by regressing $X(t)$ on $W(t)$	3.82	94.51	0.723	1.053	1.108	1.108
Calibration/season: estimate $(\widehat{a}_2, \widehat{b}_2)$ by regressing $X(t)$ on $\widehat{\kappa}_W\{W(t) - \widehat{\varsigma}(t)\}$	4.31	94.90	0.725	1.050	1.114	1.114
Calibration/season: estimate $(\widehat{a}_3, \widehat{b}_3)$ by regressing $X(t)$ on $\widehat{\varsigma}(t)(1 - \widehat{\kappa}_W) + \widehat{\kappa}_W W(t)$; see (2)	3.87	94.51	0.731	1.053	1.130	1.130
Raw data	9.14	95.13	0.886	1.028	1.665	1.664
Seasonal adjustment only	4.39	94.39	0.950	1.052	1.912	1.912
Calibration only	7.65	95.28	0.675	1.029	0.964	0.965

n is the number of case–control pairs; a is the calibration curve intercept ($a = 5$); b is the calibration curve slope ($b = 1.4$); CI% is the estimated coverage of a 95% nominal confidence interval for α_1; α_1 is the log odds ratio increase per increase in vitamin D by one standard deviation ($\tau = 15.5206$); $\sigma_{XX} = \sigma_{WW}$ is the measurement error variances ($\sigma_{WW} = 16$); 'SE' is the empirical estimate of standard error based on simulations; 'SE ratio' is the ratio of the empirical SE estimate to the mean model-based estimate of standard error; 'Var ratio' is the square of the ratio of the empirical SE for an estimator to the empirical SE of the estimate based on known exposure; 'MSE ratio' is the ratio of the mean square error from a given procedure to that based on known exposure; 'Calibration only' is defined in Section 3.1.

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

studies. All the estimates (even those based on the true values x_0) were very slightly upwardly biased, with most bias values near 4×10^{-5}, but biases were about twice this high in analyses of the true (but unknown) x_0 values, of raw vitamin D data, and of calibrated data without seasonal adjustment. All the biases were small, however, in comparison with the empirical estimates of the standard error of the estimates; see the column headed 'SE($\widehat{\alpha}_1$)'. The squared ratios of empirical standard error to the empirical standard error for the true data x_0 (see column headed 'Var ratio') were near 1.1 for each of the methods that adjusted for both seasonality and calibration, 0.964 for the method based on calibration only, 1.912 for the method that adjusted for season only, and 1.665 for the raw data. Thus, there was about a 10% efficiency loss, measured as a variance ratio, for the methods that adjust for season and calibration. The corresponding ratios of mean square errors were almost identical to the variance ratios, because the bias was negligible. All procedures yielded coverage near the nominal 95% level. However, with 10,000 simulations, the 95% confidence interval is approximately equal to the estimated coverage plus or minus 0.43%. Hence, some of the coverage rates are statistically significantly less than 95%, although only very slightly less than nominal. This probably reflects the fact that the ratios of the empirical estimates from simulations of the standard error of $\widehat{\alpha}_1$ to the average model-based estimates of standard error range from 1.028 to 1.053 (see the column headed 'SE ratio'). For example, if the ratio of the true SE to the estimated SE is 1.04, the coverage of the nominal 95% confidence interval at the null hypothesis would be $\Phi(1.96/1.04) - \Phi(-1.96/1.04) = 0.9405$, where Φ is the standard normal distribution function. Although the model-based variance estimates allow for variability in estimates of calibration parameters, they do not take variability in estimation of the trend into account. With $n = 200$ case–control pairs, this problem disappeared, as shown in Table SII, where all coverages were within 0.43% of the nominal 95%. With the larger samples, the efficiency loss from calibration and seasonal adjustment was again about 10% (Table SII).

In summary, under the null hypothesis, all procedures had nominal coverage for $n = 200$ and very near nominal coverage for $n = 100$. The efficiency loss from calibration and seasonal adjustment was about 10%, compared with the ideal situation with known x_0, but the loss in efficiency was much greater for procedures that do not calibrate the local assay data.

Table II contains comparisons under the alternative $\alpha_1 = -0.04466$, which corresponds to a halving of risk for an increase in vitamin D of $\tau = 15.52$ nmol/L, which is the estimated standard deviation of w in the ATBC data. The coverages of all procedures that adjust for season and calibration were within 0.43% of the nominal 95% level and therefore did not deviate statistically significantly from it. Analyses of the raw vitamin D data, data that are seasonally adjusted but not calibrated, and data that are calibrated but not seasonally adjusted led to coverage significantly below nominal levels. Such coverages are even smaller with $n = 200$ case–control pairs (Table SIII). Biases from failure to calibrate contributed importantly

Table II. Simulation performance of procedures for conditional logistic regression analysis of continuous vitamin D based on data from the Alpha-Tocopherol, Beta-Carotene study at the alternative hypothesis $\alpha_1 = -0.04466$, with $n = 100$ case–control pairs.

	Bias of α_1 ($\times 10^3$)	CI%	SE ($\widehat{\alpha}_1$) ($\times 10^2$)	SE ratio	Var ratio	MSE ratio
Exposure x measured without error	−173	95.11	1.028	1.052	1.000	1.000
Calibration/season: estimate $(\widehat{a}_1, \widehat{b}_1)$ by regressing $X(t)$ on $W(t)$	−153	95.06	1.084	1.061	1.112	1.103
Calibration/season: estimate $(\widehat{a}_2, \widehat{b}_2)$ by regressing $X(t)$ on $\widehat{\kappa}_W\{W(t) - \widehat{\zeta}(t)\}$	−163	95.43	1.122	1.030	1.192	1.184
Calibration/season: estimate $(\widehat{a}_3, \widehat{b}_3)$ by regressing $X(t)$ on $\widehat{\zeta}(t)(1 - \widehat{\kappa}_W) + \widehat{\kappa}_W W(t)$; see (2)	−199	95.11	1.099	1.065	1.143	1.148
Raw data	−1042	92.06	1.254	1.044	1.489	2.447
Seasonal adjustment only	−1607	83.92	1.374	1.060	1.786	4.113
Calibration only using $(\widehat{a}_1, \widehat{b}_1)$	276	90.76	0.991	1.048	0.929	0.974

Here, CI% represents the actual coverage of a 95% nominal confidence interval. Also, α_1 is the log odds ratio per increase in vitamin D by one standard deviation ($\tau = 15.5206$); $\sigma_{XX} = \sigma_{WW}$ is the measurement error variances ($\sigma_{WW} = 16$); $a = 5$; $b = 1.4$. 'SE' is the empirical estimate of standard error based on simulations. 'SE ratio' is the ratio of the empirical SE estimate to the mean model-based estimate of standard error. 'Var ratio' is the square of the ratio of the empirical SE for an estimator to the empirical SE of the estimate based on known exposure. 'MSE ratio' is the ratio of the mean square error from a given procedure to that based on known exposure.

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

to the mean squared error ratios (Tables II and SIII). With $n = 100$, all the procedures that adjusted for seasonality and calibration had comparatively small bias and losses in efficiency compared with analysis with known x_0, with variance ratios ranging from 1.112 to 1.192 (Table II). With $n = 200$, the variance ratios ranged from 1.207 to 1.385 (Table SIII).

Unreported simulations indicated similar results when σ_{WW} was misspecified as 36 when the true value was 16 and for other values of the calibration parameters.

These simulations suggest that simple procedures that adjust for seasonality and calibration will perform well in practice. In particular, the deattenuated estimates $\hat{\alpha}_1 = \hat{\alpha}_1^*/\hat{b}_1\hat{\kappa}_W$ and $\hat{\alpha}_1 = \hat{\alpha}_1^*/\hat{b}_3\hat{\kappa}_W$ had favorable bias and efficiency characteristics.

3.3. Analysis of categorical vitamin D

Recall that for $j = 1, 2, 3$, $\hat{x}_{0j} = \hat{a}_j + \hat{b}_j\hat{\varsigma}(t_0) + \hat{b}_j\hat{\kappa}_W\{W(t) - \hat{\varsigma}(t)\}$, and β is the log odds per category increase.

The cut-points used to generate the relative risks from Equation (9) were the quintiles of the control values x_0 for $t_0 = 39$ in ATBC, namely, 49.6, 58.7, 70.8, and 86.9 nmol/L (Section 2.5.2). The intercept was $\mu = -4$. Table III summarizes performance under the null hypothesis with $n = 100$ cases and matched controls. As for the continuous vitamin D analyses, we assumed realistic estimates of measurement error, $\sigma_{XX} = \sigma_{WW} = 16$, and substantial miscalibration with intercept $a = 5$ and $b = 1.4$. In Tables III and IV, the labels 'CLR using (10) with $\hat{x}_{01}$' and 'CLR using (10) with $\hat{x}_{03}$' denote procedures that substitute the calibrated and seasonally adjusted estimates of x_0, namely, $\hat{x}_{01}$ and $\hat{x}_{03}$, into the conditional likelihood Equation (10). The 'normal model' procedure used Equations (11) and (12). These three procedures that adjust for season and calibration had near nominal coverage for $\beta_2, \beta_5,$ and β (Table III) as well as for the log odds β_3 and β_4 (not shown). Adjustment for calibration alone yielded near nominal coverage. Coverage for β_5 was above nominal levels for analyses of raw values and for seasonal adjustment only.

Table III. Simulation performance of procedures for conditional logistic regression analysis of categorical vitamin D based on data from the Alpha-Tocopherol, Beta-Carotene study at the null hypothesis $\beta_i = (i-1)\beta = 0$, with $n = 100$ case–control pairs.

	Coverage (%)			Bias × 10^2			SE			MSE ratio		
	β_2	β_5	β	β_2	β_5	β	β_2	β_5	β	β_2	β_5	β
Category measured without error	95.32	94.93	95.14	−1.031	−0.665	−0.024	0.588	0.493	0.106	1	1	1
CLR using (10) with $\hat{x}_{01}$ defined using $(\hat{a}_1, \hat{b}_1)$	94.95	95.10	94.98	−10.6	−5.39	−0.687	0.631	0.522	0.111	1.285	1.133	1.352
CLR using (10) with $\hat{x}_{03}$ defined using $(\hat{a}_3, \hat{b}_3)$	95.16	95.20	95.01	−9.11	−1.38	−0.379	0.562	0.505	0.109	1.013	1.047	1.067
Normal model using (11) and (12)	94.74	95.63	95.11	−17.8	−1.76	−0.337	1.071	0.583	0.114	3.684	1.398	1.168
Raw data	95.78	100	95.96	0.444	−0.219	−0.096	0.461	1.592	0.199	0.682	10.4	3.510
Seasonal adjustment only	95.76	100	95.27	−2.74	31.5	2.178	0.391	2.073	0.157	0.807	18.1	10.9
Calibration only using $\left(\hat{a}_1, \hat{b}_1\right)$	95.65	95.92	95.32	0.692	−0.849	−0.104	1.067	0.492	0.113	0.762	0.996	0.915

Here, $(\hat{a}_1, \hat{b}_1)$ and $(\hat{a}_3, \hat{b}_3)$ are different calibration parameter estimates as described in Section 2 and Table I. Here, $a = 5$ and $b = 1.4$ are the calibration curve intercept and slope, respectively; $\sigma_{XX} = \sigma_{WW} = 16$ are the measurement error variances. 'SE' is the empirical estimate of standard error based on simulations. 'MSE ratio' is the ratio of the mean square error from a given procedure to that based on known exposure. β_2, β_5, and β are respectively the log odds ratios corresponding to the second quintile group, fifth quintile group, and per quintile category increase (trend). The reference date was $t_0 = 39/52$ (week 39). The 20th, 40th, 60th, and 80th percentiles of control measurements adjusted to t_0 were 49.61, 58.73, 70.80, and 86.89 nmol/L, respectively.

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

Large bias was seen for seasonal adjustment alone for β and β_5. Standard errors for $\widehat{\beta}_5$ and $\widehat{\beta}$ were larger for the raw data analysis and for seasonal adjustment alone than for CLR using (10) with $\widehat{x}_{01}$ or $\widehat{x}_{03}$. The mean square error ratios (compared with the analysis with perfect x_0 data) indicated poor performance for raw data and seasonal adjustment only. For estimating β_2, the normal model had higher mean square error than CLR using (10) with $\widehat{x}_{01}$ or $\widehat{x}_{03}$, primarily because it had larger standard error. The higher variability of $\widehat{\beta}_2$ from the normal model is not the result of estimating trend parameters, calibration parameters, or κ_W, as shown in unreported simulations using known values for these parameters. Similar results were obtained with $n = 200$ cases and controls (Table SIV), except that subnominal coverage of β_2 was found for the raw data and seasonal adjustment only. These results were obtained with the reference date $t_0 = 39$ weeks. Similar results were obtained with $t_0 = 1$, in which quintiles were re-estimated to correspond to the new reference date (unreported data). Results were little affected if the variances for measurement error were mistakenly set at $\sigma_{XX} = \sigma_{WW} = 36$, when in fact $\sigma_{XX} = \sigma_{WW} = 16$ (unreported data).

Under the alternative $\beta_i = (i-1)\beta$, with $\beta = -0.25\log(2) = -0.17329$, similar results were found. With $n = 100$ cases and controls (Table IV), the coverage was near the nominal 95% level except for some analyses with raw data and with seasonal adjustment only. Bias was comparatively large for analysis of the raw data, seasonal adjustment only, and the normal model. Standard errors were comparatively large for the normal model. The mean square error ratios were large for raw data analysis, seasonal adjustment only, and the normal model for $\widehat{\beta}_2$. Similar results were found for $n = 200$ cases and controls (Table SV). Results were little affected by misspecifying $\sigma_{XX} = \sigma_{WW} = 36$, when in fact $\sigma_{XX} = \sigma_{WW} = 16$ (unreported data).

Additional simulations based on PLCO data support these findings for continuous and categorical analyses and are in Tables SVI–SIX.

To summarize, CLR using (10) with $\widehat{x}_{01}$ and CLR using (10) with $\widehat{x}_{03}$ outperformed other procedures. CLR using (10) with $\widehat{x}_{03}$ usually had smaller MSE than CLR using (10) with $\widehat{x}_{01}$, however, not only under

Table IV. Simulation performance of procedures for conditional logistic regression analysis of categorical vitamin D based on data from the Alpha-Tocopherol, Beta-Carotene study at the alternative hypothesis $\beta_i = (i-1)\beta$, with $\beta = -0.25\log(2) = -0.17329$ and with $n = 100$ case–control pairs.

	Coverage (%)			Bias × 10 ^{2}			SE			MSE ratio		
	β_2	β_5	β	β_2	β_5	β	β_2	β_5	β	β_2	β_5	β
Category measured without error	95.63	95.20	95.33	−0.921	−3.60	−0.347	0.524	0.507	0.107	1	1	1
CLR using (10) with $\widehat{x}_{01}$ defined using $(\widehat{a}_1, \widehat{b}_1)$	94.95	95.10	94.98	−2.76	−1.21	−0.083	0.595	0.522	0.111	1.282	1.075	0.965
CLR using (10) with $\widehat{x}_{03}$ defined using $(\widehat{a}_3, \widehat{b}_3)$	94.97	94.95	95.29	−5.47	2.55	0.528	0.533	0.505	0.109	1.032	1.041	0.977
Normal model using (11) and (12)	94.75	95.33	95.45	−16.0	2.95	−0.337	1.014	0.599	0.114	3.727	1.396	1.080
Raw data	94.86	100	95.64	−18.4	53.3	−4.02	0.492	1.518	0.226	1.901	10.0	26.2
Seasonal adjustment only	94.49	100	94.90	−18.3	67.2	−1.06	0.406	2.014	0.173	2.230	17.5	37.5
Calibration only using $\left(\widehat{a}_1, \widehat{b}_1\right)$	95.48	94.66	94.04	−4.94	8.05	−0.506	0.477	0.534	0.105	0.847	1.131	1.364

Here, $(\widehat{a}_1, \widehat{b}_1)$ and $(\widehat{a}_3, \widehat{b}_3)$ are different calibration parameter estimates as described in Section 2 and in Table I. Here, $a = 5$ and $b = 1.4$ are the calibration curve intercept and slope, respectively; $\sigma_{XX} = \sigma_{WW} = 16$ are the measurement error variances. 'SE' is the empirical estimate of standard error based on simulations. 'MSE ratio' is the ratio of the mean square error from a given procedure to that based on known exposure. β_2, β_5, and β are respectively the log odds ratios corresponding to the second quintile group, fifth quintile group, and per quintile category increase (trend). The reference date was $t_0 = 39/52$ (week 39). The 20th, 40th, 60th, and 80th percentiles of control measurements adjusted to t_0 were 49.61, 58.73, 70.80, and 86.89 nmol/L, respectively.

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

Table V. Log odds ratios for the ATBC Colorectal Cancer Study with Institute of Medicine cut-points and seasonal adjustment to week 39 and to week 1.

	Adjusted for season and calibration							
	CLR using (10) with $\hat{x}_{01}$	$\hat{\alpha}_1^*/\hat{b}_1\hat{\kappa}_W$	CLR using (10) with $\hat{x}_{03}$	$\hat{\alpha}_1^*/\hat{b}_3\hat{\kappa}_W$	Normal model	No adjustment (raw)	Season only	Calibration only
				Reference week 39				
Q_2	−10.9 (78.2)		−1.81* (0.684)		−4.53* (1.16)	0.138 (0.237)	−0.749* (0.276)	−0.653* (0.302)
Q_3	−10.8 (78.2)		−1.54* (0.674)		−3.89* (1.05)	−0.510 (0.331)	−0.950* (0.331)	−0.572 (0.330)
Q_4	−11.1 (78.2)		−1.97* (0.712)		−4.46* (1.11)	−0.603 (0.671)	−1.20* (0.572)	−1.36* (0.431)
CAT	−0.219 (0.142)		−0.175 (0.137)		−0.308* (0.148)	−0.177 (0.133)	−0.418* (0.144)	−0.334* (0.126)
CONT		−0.143* (0.061)		−0.143* (0.061)		−0.168* (0.070)	−0.171* (0.070)	−0.140* (0.061)
				Reference week 1				
Q_2	−0.583* (0.278)		−0.300 (0.249)		−0.415 (0.330)	0.138 (0.237)	0.166 (0.233)	−0.653* (0.302)
Q_3	−0.586 (0.305)		−0.457 (0.304)		−0.399 (0.342)	−0.510 (0.331)	−0.303 (0.334)	−0.573 (0.330)
Q_4	−1.53* (0.505)		−1.25* (0.526)		−1.29* (0.570)	−0.603 (0.671)	−0.223 (0.846)	−1.36* (0.431)
CAT	−0.366* (0.127)		−0.311* (0.125)		−0.300* (0.133)	−0.117 (0.133)	−0.074 (0.139)	−0.334* (0.126)
CONT		−0.143* (0.061)		−0.143* (0.061)		−0.168* (0.070)	−0.171* (0.070)	−0.140* (0.061)

Standard errors are in parentheses. 'NORMAL' indicates the normal model procedure in Equations (11) and (12). 'CLR using (10) with $\hat{x}_{01}$' and 'CLR using (10) with $\hat{x}_{03}$' denote procedures that substitute those calibrated and seasonally adjusted estimates of x_0 into the conditional logistic likelihood (10). The quantities Q_2, Q_3, and Q_4 denote log odds ratios for categories 2, 3, and 4, respectively, compared with category 1. 'CAT' indicates the log odds per category increase in a trend model with scores 0, 1, 2, and 3. 'CONT' refers to the continuous trend model with the trend parameter being the log odds per 10 nmol/L increase in vitamin D. Estimates of $\tau, a, b, \gamma_0, \gamma_1, \gamma_2, \gamma_3,$ and γ_4 from the ATBC data were 15.52, 3.86, 1.20, 39.43, −5.84, −4.49, 1.44, and −0.444, respectively.

*$P < 0.05$.

the null hypothesis (Tables III, SIV, and SVIII) but also under the alternative (Tables IV, SV, and SIX). Thus, we recommend CLR using (10) with $\hat{x}_{03}$.

4. Examples of analyses of three studies

We present analyses for the colorectal case–control study in ATBC [3] and for two case–control studies of breast cancer, the NYUWHS [9] and a study from the CPSII Nutrition Cohort [10]. Table V presents categorical analyses for the ATBC data compared with the Institute of Medicine cut-points (<30, [30, 50), [50, 75), ⩾75 nmol/L), as well as analyses of continuous vitamin D as a linear effect. The categories correspond respectively to vitamin D deficiency, probable insufficiency, adequacy, and no increased benefit. The ATBC study included 146 colorectal cancer cases and 290 matched controls. Estimates $\hat{\tau}, \hat{a}_1$, and $\hat{b}_1$ and estimates of the seasonal trend parameters are in the footnote of Table V. Conditional likelihoods (6), (10), and (12) were modified to accommodate varying numbers of matched controls per case. We analyzed the data with two different reference dates for seasonal adjustment to assess the robustness of the conclusions. With week 39 as the reference date, the preferred analyses of the categorical data based on CLR using (10) with $\hat{x}_{03}$ indicated a strong protective effect of increased vitamin D with a statistically significant odds ratio of $\exp(-1.97) = 0.139$ for the highest category (⩾75 nmol/L). Indeed, the log odds ratios for all categories above the reference category (<30 nmol/L) were statistically significantly less than 0; the categorical trend with odds per category increase of $\exp(-0.175) = 0.839$ was not statistically significant. Analyses based on CLR using (10) with $\hat{x}_{01}$ were not stable because only six cases and two controls were categorized below 30 nmol/L. Note that analyses without any adjustment, with seasonal adjustment alone and with calibration alone also indicated protective trends, but the parameter estimates were smaller and usually not statistically significantly different from 0 in the absence of seasonal adjustment. The continuous trend analysis, which does not depend on a reference date, indicated a statistically significant protective odds ratio of $\exp(-0.143) = 0.866$ per 10 nmol/L increase in vitamin D concentration. When these data were re-analyzed using week 1 as the reference date, the magnitudes of the protective effects in categorical analyses were smaller, but the qualitative conclusion of a significant protective effect remained unchanged. For example, the estimated protective odds ratio for vitamin D ⩾75 nmol/L was $\exp(-1.25) = 0.287$, instead of the previous 0.139. Analyses based on a linear trend in continuous vitamin D were unchanged.

Categorical analyses of the NYUWHS breast cancer data (893 cases and 1642 controls) with a reference date of week 39 suggested a slight favorable effect from increasing vitamin D levels, in line with the analysis based on a linear trend in continuous vitamin D (Table SX). Re-analysis with week 1 as the reference date yielded similar small favorable categorical effects that were not statistically significant. Analyses of continuous vitamin D were unchanged.

We performed similar analyses for breast cancer data from CPSII (515 cases and 515 controls) (Table SXI). With week 39 as the reference date, categorical analyses suggested a slight protective effect of increasing vitamin D, whereas the continuous analysis indicated a slight adverse effect. None of these effects were statistically significant, however. With week 1 as the reference date, categorical effects were less protective. However, all these estimates were consistent with the null hypothesis of no vitamin D effect.

These examples illustrate that for externally selected cut-points, such as the Institute of Medicine guidelines, estimates of category-specific log odds ratios are affected by the choice of reference date for seasonal adjustment, unlike analyses of a continuous linear trend.

5. Discussion

We presented a framework for seasonal adjustment and calibration of matched case–control data and used this framework to evaluate analytic procedures for continuous and categorical vitamin D risk models. We evaluated the procedures though simulations based on the Vitamin D Pooling Project data. Several practical conclusions and recommendations follow from this work: (1) adjustment for calibration and seasonal trends yields more nearly nominal coverage of confidence intervals, less bias, smaller variance, and reduced mean square error than unadjusted analyses; (2) simple procedures are available to adjust for calibration and seasonal trend; (3) for calibration samples obtained by stratified random sampling, as in the Vitamin D Pooling Project, there is surprisingly little inflation in the variance of estimates of vitamin D effects, even with only 30 calibration measurements; and (4) estimates for the logit of risk for

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

continuous vitamin D are independent of a reference date for seasonal adjustment. In contrast, estimates for a categorical risk model depend on the reference date. Thus, public health recommendations based on vitamin D categories may need to be season specific. We elaborate on some of these points.

Simple procedures are available for seasonal adjustment and calibration to estimate the slope, α_1, of a linear trend in vitamin D on the logit of risk. First, one estimates the slope α_1^* by using seasonally detrended values $\{W(t) - \widehat{\varsigma}(t)\}$ in place of vitamin D measurements in CLR. Then, one deattenuates and corrects for calibration by setting $\widehat{\alpha}_1 = \widehat{\alpha}_1^* / \widehat{b}_3 \widehat{\kappa}_W$. Our methods do not adjust variances for estimating seasonal trend parameters. Nonetheless, even with $n = 100$ cases and controls, estimating the five seasonal trend parameters had little impact on the coverage of confidence intervals for α_1, and with $n = 200$, there was no evidence of impact. In Supporting Information Appendix A.6, we justify this result asymptotically. There is some loss of efficiency in estimating α_1 from estimating the calibration curve, but the additional variance is accounted for in constructing confidence intervals (Section 2.5.1).

For categorical data, CLR (10) with the estimate $\widehat{x}_{03}$ can be recommended. In this procedure, the calibration estimates obtained by regressing $X(t)$ on $\widehat{\varsigma}(t)(1 - \widehat{\kappa}_W) + \widehat{\kappa}_W W(t)$ are used, together with the seasonal trend estimate, to compute $\widehat{x}_{03}$ from Equation (7). Then, $\widehat{x}_{03}$ is categorized as if it were x_0, and the analysis proceeds using standard CLR. No special calculations were used to accommodate variation from estimating seasonal trend or calibration parameters. This procedure with $\widehat{x}_{01}$ in place of $\widehat{x}_{03}$ worked nearly as well but is less theoretically justified because it does not take seasonal adjustment and $\widehat{\kappa}_W$ into account (Equation (2)) and usually has larger mean square error. For $\widehat{\kappa}_W = 1$, however, the two procedures are equivalent. The normal model method has theoretical appeal but tended to yield more variable estimates of log odds parameters, especially β_2. This loss of efficiency seems inherent in the likelihood based on Equation (12). Unreported simulations show that even if the parameters a, b, and κ_W are known, the normal model method yields more variable estimates than the procedures that use $\widehat{x}_{01}$ or $\widehat{x}_{03}$ in Equation (10).

The sine–cosine series in Equation (4) fit the data well, required only five parameters, and forced the estimated trend to be periodic. Flexible procedures such as LOESS do not yield periodic trends, are more sensitive to gaps in the data, and can require many parameters, leading to increased variance of estimates of vitamin D effects [7].

We assumed normal distributions in our analytical framework (Section 2.2). More work is needed to assess the performance of our recommended procedures for other distributions.

An important practical issue was brought out by the examples in Section 4. The Institute of Medicine guidelines for categories of vitamin D levels do not specify the time of year the blood sample was drawn. In the examples, vitamin D levels fluctuated by about 16 nmol/L between the highest values, which occurred in September–October, and the lowest values, which occurred in January–April. Thus, a person might be categorized in the highest Institute of Medicine category in September but lower in February. The intent of seasonal adjustment is to transform each person's vitamin D value to the value that person would have had if measured on a common reference date, such as week 39 or week 1. This improves comparability between cases and controls, especially in studies without matching on date of blood draw. The odds ratios depend on the choice of reference date (Appendix A.2). Thus one should specify the reference date used for categorical analyses applied to fixed cut-points. If, instead of using externally determined cut-points such as the Institute of Medicine categories, one standardizes the control data to $\{W(t) - \widehat{\varsigma}(t) + \widehat{\varsigma}(t_0)\}$ and chooses study-specific cut-points as percentiles of the standardized control data, then the categorical analysis is much less sensitive to choice of reference date, t_0, as indicated by unreported simulations. Analysis of continuous vitamin D as a linear trend is not affected by choice of reference date.

An alternative to picking a particular reference date, t_0, is to average the seasonal estimate $\widehat{\varsigma}(t_0)$ over the 52 weeks of the year. Because integrated sinusoidal terms vanish over the year, this is equivalent to replacing $\widehat{\varsigma}(t_0)$ by $\widehat{\gamma}_0$, the estimated intercept in Equation (4). The resulting log odds estimates for Q_2, Q_3, and Q_4 for the data in Table V from Equation (10) with $\widehat{x}_{03}$ were -1.176, -1.160, and -1.582, respectively, and are intermediate between the values shown in Table V for reference weeks $t_0 = 39$ (high vitamin D levels) and $t_0 = 1$ (low vitamin D levels).

To summarize, simple methods for calibration and seasonal adjustment yield reliable inference for matched case–control data on vitamin D, with only modest loss of statistical efficiency, compared with analyses with perfectly known vitamin D. These findings are informing analytic approaches for the Vitamin D Pooling Project.

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

Appendix A

A.1. Justification for the normal model Equation (12)

Let the categories be $(C_1, \ldots, C_5)$, and write the risk model (9) as

$$\text{pr}(Y = 1|x) = H\left(\sum_{j=1}^{5} I(x \in C_j)\beta_j\right) = \sum_{j=1}^{5} H(\beta_j) I(x \in C_j),$$

where $H(u) = \{1 + \exp(-u)\}^{-1}$ is the logistic function. Assuming that Y is conditionally independent of $W(t)$ given that x is in category c,

$$\text{pr}\{Y = 1|W(t)\} = \sum_{j=1}^{5} H(\beta_j)\text{pr}\{x \in C_j|W(t)\}.$$

Consider the matched pairs problem, with the matching effect α_i belonging to the ith matched pair, with responses (Y_{i1}, Y_{i2}), predictors (x_{i1}, x_{i2}), and $\{W_{i1}(t), W_{i2}(t)\}$. We have that

$$\begin{aligned}
\text{pr}(Y_{ik} &= 1|x_{ik}) = \sum_{j=1}^{5} H(\alpha_i + \beta_j) I(x_{ik} \in C_j); \\
\text{pr}\{Y_{ik} &= 1|W_{ik}(t)\} = \sum_{j=1}^{5} H(\alpha_i + \beta_j)\text{pr}\{x_{ik} \in C_j|W_{ik}(t)\}.
\end{aligned}$$

Under the rare disease approximation $H(\alpha_i + \beta_j) \approx \exp(\alpha_i + \beta_j)$. It follows that

$$\begin{aligned}
&\text{pr}\{Y_{i1} = 1, Y_{i2} = 0|W_{i1}(t), W_{i2}(t), Y_{i1} + Y_{i2} = 1\} \\
&\quad \approx \frac{\sum_{j=1}^{5} \exp(\beta_j)\text{pr}\left\{(x_{i1} \in C_j|W_{i1}(t)\right\}}{\sum_{j=1}^{5} \exp(\beta_j)\text{pr}\left\{(x_{i1} \in C_j|W_{i1}(t)\right\} + \sum_{j=1}^{5} \exp(\beta_j)\text{pr}\{x_{i2} \in C_j|W_{i2}(t)\}},
\end{aligned}$$

which equals Equation (12).

A.2. Impact of seasonal trend adjustment, miscalibration, and variation in cut-points on categorical odds ratios

If fixed cut-points are used to define categories of exposure, then a change in the cut-points, a change in the reference date for seasonal adjustment, or miscalibration will induce bias in odds ratios with respect to the original model.

Changing the cut-points

For simplicity, we assume only two categories, but the arguments extend to multiple categories. We assume initially that the true exposure at reference date t_0, namely, $x_0 = x + \varsigma(t_0)$, is known (no measurement error and no miscalibration). Suppose that $x_0 > c$ defines the high-risk group and $x_0 \leqslant c$ the low-risk group. Let p_1 be the risk of disease in the high-risk group and p_0 be the risk of disease in the low-risk group. The odds ratio is $OR = \{p_1/(1 - p_1)\}/\{p_0/(1 - p_0)\}$. Now, suppose we choose another cut-point c^*. Let F be the distribution of x_0 in the population and f be the density of x_0. Suppose that the risk in the population is determined by the model

$$P(x_0, Y = 1) = f(x_0)\{p_1 I(x_0 > c) + p_0 I(x_0 \leqslant c)\}, \tag{A.1}$$

where $Y = 1$ indicates diseased. The portion in curly brackets is a special case of Equation (9) of the paper. If we use a new cut-point, c^*, and assume $c^* > c$,

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

$$
\begin{aligned}
OR^* &= \{P(x_0 > c^*, Y = 1)/P(x_0 > c^*, Y = 0)\}/\{P(x_0 \leqslant c^*, Y = 1)/P(x_0 \leqslant c^*, Y = 0)\} \\
&= \frac{[\{P(x_0 \leqslant c, Y = 0) + P(c < x_0 \leqslant c^*, Y = 0)\}\{1 - F(c^*)\}p_1/\{1 - F(c^*)\}(1 - p_1)]}{\{P(x_0 \leqslant c, Y = 1) + P(c < x_0 \leqslant c^*, Y = 1)\}} \\
&= \frac{[\{1 - F(c^*)\}p_1/\{1 - F(c^*)\}(1 - p_1)]}{[\{F(c)p_0 + \{F(c^*) - F(c)\}p_1)]/[F(c)(1 - p_0) + \{F(c^*) - F(c)\}(1 - p_1)]} \\
&= \frac{[p_1/(1 - p_1)]}{[\{F(c)p_0 + \{F(c^*) - F(c)\}p_1)]/[F(c)(1 - p_0) + \{F(c^*) - F(c)\}(1 - p_1)]}.
\end{aligned}
\tag{A.2}
$$

For example, if F is a standard normal distribution, $c = 0$, $p_1 = 0.05$, and $p_0 = 0.01$, then $OR = 5.21$. If, instead, $c^* = 1.645$, the 95th percentile of F, then

$$
OR^* = (0.05/0.95)/(0.5 \times 0.01 + 0.45 \times 0.05/0.5 \times 0.99 + 0.45 \times 0.95) = 0.05263/0.0298 = 1.765,
$$

which is severely downwardly biased. Thus, changing the cut-point changes the odds ratios if Equation (A.1) is the correct model. A similar argument can be used to compute OR^* if $c^* \leqslant c$ (Equation (A.3)).

Seasonal adjustment

Suppose we chose a different date for seasonal adjustment, t_0^*. This would lead to $x_0^* = x + \varsigma(t_0^*) > c$,which is equivalent to $x_0 > c - \varsigma(t_0^*) + \varsigma(t_0) \equiv c^*$. Thus, assuming the model based on t_0 is correct, we can compute the bias from using t_0^* by substituting c^* in Equation (A.2), provided $c^* > c$. If $c^* \leqslant c$,

$$
\begin{aligned}
OR^* &= \{P(x_0 > c^*, Y = 1)/P(x_0 > c^*, Y = 0)\}/\{P(x_0 \leqslant c^*, Y = 1)/P(x_0 \leqslant c^*, Y = 0)\} \\
&= \frac{\{P(c^* < x_0 \leqslant c, Y = 1) + P(c < x_0, Y = 1)\}/\{P(c^* < x_0 \leqslant c, Y = 0) + P(c < x_0, Y = 0)\}}{\{P(x_0 \leqslant c^*, Y = 1)\}/\{P(x_0 \leqslant c^*, Y = 0)\}} \\
&= \frac{[\{F(c) - F(c^*)\}p_0 + \{1 - F(c)\}p_1]/[\{F(c) - F(c^*)\}(1 - p_0) + \{1 - F(c)\}(1 - p_1)]}{\{p_0/(1 - p_0)\}}.
\end{aligned}
\tag{A.3}
$$

Miscalibration

Suppose we use a miscalibrated variable $z_0 = a + bx_0$ instead of x_0 but retain the cut-point c for x_0. Then, if $b > 0$, the event $z > c$ is equivalent to $x_0 > (c - a)/b \equiv c^*$. Hence, the biased OR^* can be computed from Equation (A.2) or (A.3) accordingly as $c^* > c$ or $c^* \leqslant c$. If $b < 0$, then $z > c$ is equivalent to $x_0 < (c - a)/|b| \equiv c^*$. If $c^* < c$ and F is continuous at cut-points, then OR^* can be calculated from

$$
\begin{aligned}
OR^* &= \{P(x_0 < c^*, Y = 1)/P(x_0 < c^*, Y = 0)\}/\{P(x_0 \geqslant c^*, Y = 1)/P(x_0 \geqslant c^*, Y = 0)\} \\
&= \frac{F(c^*)p_0/F(c^*)(1 - p_0)}{\{P(x_0 \geqslant c, Y = 1) + P(c^* \leqslant x_0 < c, Y = 1)\}/\{P(x_0 \geqslant c, Y = 0) + P(c^* \leqslant x_0 < c, Y = 0)\}} \\
&= \frac{[\{1 - F(c^*)\}p_1/\{1 - F(c^*)\}(1 - p_1)]}{[\{F(c)p_0 + \{F(c^*) - F(c)\}p_1)]/[F(c)(1 - p_0) + \{F(c^*) - F(c)\}(1 - p_1)]} \\
&= \frac{[p_0/(1 - p_0)]}{[\{1 - F(c)\}p_1 + \{F(c) - F(c^*)\}p_0)]/[\{1 - F(c)\}(1 - p_1) + \{F(c) - F(c^*)\}(1 - p_0)]}.
\end{aligned}
\tag{A.4}
$$

If $c^* \geqslant c$,

$$
\begin{aligned}
OR^* &= \{P(x_0 < c^*, Y = 1)/P(x_0 < c^*, Y = 0)\}/\{P(x_0 \geqslant c^*, Y = 1)/P(x_0 \geqslant c^*, Y = 0)\} \\
&= \frac{\{P(x_0 < c, Y = 1) + P(c \leqslant x_0 < c^*, Y = 1)\}/\{P(x_0 < c, Y = 0) + P(c \leqslant x_0 < c^*, Y = 0)\}}{P(x_0 \geqslant c^*, Y = 1)/\{P(x_0 \geqslant c^*, Y = 0)} \\
&= \frac{[F(c)p_0 + \{F(c^*) - F(c)\}p_1)]/[\{F(c)\}(1 - p_0) + \{F(c^*) - F(c)\}(1 - p_1)]}{p_1/(1 - p_1)}.
\end{aligned}
\tag{A.5}
$$

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* **2016,** 35 2133–2148

Miscalibration and seasonal adjustment

If $z_0^* = a + bx_0^* = a + b\{x + \varsigma(t_0^*)\} = a + b\{x_0 - \varsigma(t_0) + \varsigma(t_0^*)\}$, then $z_0^* > c$ is equivalent to $x_0 > \{(c-a)/b\} + \varsigma(t_0) - \varsigma(t_0^*) \equiv c^*$ provided $b > 0$. Hence, OR^* can be computed from (A.2) or (A.3) accordingly as $c^* > c$ or $c^* \leqslant c$. If $b < 0$, the condition is $x_0 < \{(c-a)/|b|\} + \varsigma(t_0) - \varsigma(t_0^*) \equiv c^*$. Then, OR^* can be computed from (A.4) or (A.5) accordingly as $c^* < c$ or $c^* \geqslant c$.

Summary

These calculations indicate that for categorical analyses the definition of the odds ratio of interest depends on the reference date for seasonal adjustment. They also indicate that miscalibration will affect the categorical odds ratio.

Acknowledgements

This work was supported by the Intramural Research Program of the National Cancer Institute, Division of Cancer Epidemiology and Genetics. Carroll's research was supported by a grant from the National Cancer Institute (U01-CA057030). We thank Stephanie Weinstein, Demetrius Albanes, and Michal Freedman for the use of ATBC and PLCO data; the American Cancer Society for use of the CPSII data; Li Cheung for help with R programming; and Mattias Johansson and David Muller for assistance with the sine–cosine method for seasonal adjustment. The American Cancer Society funded the creation, maintenance, and update of the CPSII cohort. The NYUWHS is supported by the National Cancer Institute (grants R01-CA098661, UM1-CA182934, and P30-CA016087) and by the National Institute of Environmental Health Sciences (center grant ES000260). The main Vitamin D Pooling Project was supported by R01-CA152071.

References

1. Munger KL, Levin LI, Hollis BW, Howard NS, Ascherio A. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. *Journal of the American Medical Association* 2006; **296**(23):2832–2838.
2. Jenab M, Bueno-de Mesquita HB, Ferrari P, van Duijnhoven FJ, Norat T, Pischon T, Jansen EH, Slimani N, Byrnes G, Rinaldi S, Tjønneland A, Olsen A, Overvad K, Boutron-Ruault MC, Clavel-Chapelon F, Morois S, Kaaks R, Linseisen J, Boeing H, Bergmann MM, Trichopoulou A, Misirli G, Trichopoulos D, Berrino F, Vineis P, Panico S, Palli D, Tumino R, Ros MM, van Gils CH, Peeters PH, Brustad M, Lund E, Tormo MJ, Ardanaz E, Rodríguez L, Sánchez MJ, Dorronsoro M, Gonzalez CA, Hallmans G, Palmqvist R, Roddam A, Key TJ, Khaw KT, Autier P, Hainaut P, Riboli E. Association between pre-diagnostic circulating vitamin D concentration and risk of colorectal cancer in European populations: a nested case–control study. *British Medical Journal* 2010; **340**:b5500.
3. Tangrea J, Helzlsouer K, Pietinen P, Taylor P, Hollis B, Virtamo J, Albanes D. Serum levels of vitamin D metabolites and the subsequent risk of colon and rectal cancer in Finnish men. *Cancer Causes & Control* 1997; **8**(4):615–625.
4. Freedman DM, Chang SC, Falk RT, Purdue MP, Huang WY, McCarty CA, Hollis BW, Graubard BI, Berg CD, Ziegler RG. Serum levels of vitamin D metabolites and breast cancer risk in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *Cancer Epidemiology Biomarkers & Prevention* 2008; **17**(4):889–894.
5. Bliss CI. Periodic regression in biology and climatology. *The Connecticut Agricultural Experiment Station Bulletin* 1958: 1–56.
6. Bolland MJ, Grey AB, Ames RW, Mason BH, Horne AM, Gamble GD, Reid IR. The effects of seasonal variation of 25-hydroxyvitamin D and fat mass on a diagnosis of vitamin D sufficiency. *American Journal of Clinical Nutrition* 2007; **86**(4):959–964.
7. Zhang H, Ahn J, Yu K. Comparing statistical methods for removing seasonal variation from vitamin D measurements in case–control studies. *Statistics and its Interface* 2011; **4**(1):85–93.
8. Institute of Medicine (US) Committee to Review Dietary Reference Intakes for Vitamin D and Calcium. In *Dietary Reference Intakes for Calcium and Vitamin D*, Ross AC, Taylor CL, Yaktine AL, Del Valle HB (eds). National Academies Press: Washington, DC, 2011.
9. Scarmo S, Afanasyeva Y, Lenner P, Koenig KL, Horst RL, Clendenen TV, Arslan AA, Chen Y, Hallmans G, Lundin E, Rinaldi S, Toniolo P, Shore RE, Zeleniuch-Jacquotte A. Circulating levels of 25-hydroxyvitamin D and risk of breast cancer: a nested case–control study. *Breast Cancer Research* 2013; **15**(1):R15.
10. McCullough ML, Stevens VL, Patel R, Jacobs EJ, Bain EB, Horst RL, Gapstur SM, Thun MJ, Calle EE. Serum 25-hydroxyvitamin D concentrations and postmenopausal breast cancer risk: a nested case control study in the Cancer Prevention Study-II Nutrition Cohort. *Breast Cancer Research* 2009; **11**(4):R64.

Supporting information

Additional supporting information may be found in the online version of this article at the publishers web-site. This material includes Tables SI–SXI and Figure S1, as well as the asymptotic theory for the continuous risk model.

Published 2016. This article has been contributed to by US Government employees and their work is in the public domain in the USA. *Statist. Med.* ***2016,*** 35 2133–2148