Title

Accommodating Measurements Below a Limit of Detection: A Novel Application of Cox Regression

Journal

American Journal of Epidemiology, 179(8)

ISSN

0002-9262

Authors

Dinse, Gregg E
Jusko, Todd A
Ho, Lindsey A
et al.

Publication Date

2014-04-15

DOI

10.1093/aje/kwu017

Peer reviewed

American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2014. This work is written by (a) US Government employee(s) and is in the public domain in the US.

Vol. 179, No. 8
DOI: 10.1093/aje/kwu017
Advance Access publication:
March 4, 2014

Practice of Epidemiology

Accommodating Measurements Below a Limit of Detection: A Novel Application of Cox Regression

Gregg E. Dinse, Todd A. Jusko, Lindsey A. Ho, Kaushik Annam, Barry I. Graubard, Irva Hertz-Picciotto, Frederick W. Miller, Brenda W. Gillespie, and Clarice R. Weinberg*

* Correspondence to Dr. Clarice R. Weinberg, Mail Drop A3-03, P.O. Box 12233, Research Triangle Park, NC 27709 (e-mail: weinber2@niehs.nih.gov).

Initially submitted April 28, 2013; accepted for publication January 15, 2014.

In environmental epidemiology, measurements of exposure biomarkers often fall below the assay's limit of detection. Existing methods for handling this problem, including deletion, substitution, parametric regression, and multiple imputation, can perform poorly if the proportion of "nondetects" is high or parametric models are misspecified. We propose an approach that treats the measured analyte as the modeled outcome, implying a role reversal when the analyte is a putative cause of a health outcome. Following a scale reversal as well, our approach uses Cox regression to model the analyte, with confounder adjustment. The method makes full use of quantifiable analyte measures, while appropriately treating nondetects as censored. Under the proportional hazards assumption, the hazard ratio for a binary health outcome is interpretable as an adjusted odds ratio: the odds for the outcome at any particular analyte concentration divided by the odds given a lower concentration. Our approach is broadly applicable to cohort studies, case-control studies (frequency matched or not), and cross-sectional studies conducted to identify determinants of exposure. We illustrate the method with cross-sectional survey data to assess sex as a determinant of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin concentration and with prospective cohort data to assess the association between 2,4,4′-trichlorobiphenyl exposure and psychomotor development.

2,3,7,8-tetrachlorodibenzo-*p*-dioxin; 2,4,4′-trichlorobiphenyl; hazard identification; limit of detection; National Health and Nutrition Examination Survey; nondetects; proportional hazards

Abbreviations: LOD, limit of detection; NHANES, National Health and Nutrition Examination Survey; PCB, polychlorinated biphenyl; PCB-28, 2,4,4′-trichlorobiphenyl; PDI, psychomotor development index; TCDD, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin.

The increasing availability of informative exposure biomarkers presents both opportunities and challenges for environmental epidemiologists as we try to identify determinants of exposure and assess the health effects of environmental pollutants. Many exposure biomarkers have low concentrations in serum, urine, or other biological matrices, with measurements often falling below the assay's limit of detection (LOD), the lowest level at which a substance's presence is distinguishable from its absence (1). Here we propose a new approach for managing these "nondetects," verify its appropriate confidence interval coverage under the null through simulations, and illustrate its use through application to 2 data examples.

Existing methods for handling nondetects include discarding them, replacing them, or multiply imputing them. The deletion method is inefficient (2) and prone to bias; in our example involving 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) in the National Health and Nutrition Examination Survey (NHANES), the deletion approach would discard 82% of the participants. Replacing nondetects with a specific value such as LOD / 2 or LOD / $\sqrt{2}$ can produce bias and overly narrow confidence intervals, as can model misspecification in parametric substitution schemes based on conditional expectations; distribution-free substitution of the average detectable value can perform poorly as well (2). Multiple imputation methods can also produce bias and overly narrow confidence intervals when the assumed parametric model is misspecified, especially in applications with many measurements below the LOD (3). In view of these difficulties, we sought a confounder-adjusted approach that would avoid

restrictive parametric assumptions while accommodating a high proportion of nondetects.

Our method treats the biomarker as the outcome, which sometimes implies switching the roles of exposure and outcome. We use Cox regression (4) to model analyte concentration as a function of the health measure, while adjusting for confounders. Without assuming a parametric concentration distribution, Cox regression easily accommodates nondetects by appropriately treating them as censored. We reverse the analyte's scale to transform the provided left-censored observations into more easily handled right-censored observations. Although Cox's proportional hazards model was developed for censored event-time data, it can be applied to any variable that is subject to (potentially heavy) censoring. In addition to adjusting for confounders, the Cox model does not require event times (or analyte concentrations in this setting) to follow any particular parametric distribution.

Role reversal of an exposure and its effect raises questions about the use of directed acyclic graphs (5) to identify possible confounders for inclusion as covariates. However, the roles of the analyte and the health measure are typically symmetrical (i.e., confounder selection is invariant to flipping the direction of that causal arrow), so that the same confounders will usually be appropriate to include as covariates. (The approach we propose would not be useful, however, for inference involving nonconfounder intermediates.)

This work was originally motivated by a study of xenobiotic exposures measured in NHANES from 1999 to 2004 (6). One chemical measured was TCDD, and 82% of its concentrations were below the LOD. Using the same NHANES data to assess associations between polychlorinated biphenyls (PCBs) and antinuclear antibodies, Gallagher et al. (7) restricted their logistic analyses to the last 2 years (2003–2004), excluding the first 4 years (1999–2002) because the less sensitive assay used earlier had generated a high proportion of nondetects. Our second example involves data on infant development in relation to in utero 2,4,4′-trichlorobiphenyl (PCB-28) exposure, where a previous analysis omitted several measured chemicals because too many values were below the LOD (8). We wanted a method that could use all of the available data. Furthermore, we wanted a method that could be applied in a wide range of settings, including case-control studies (with or without frequency matching), randomized controlled trials, and cross-sectional studies, and for which the health outcome could be dichotomous, multilevel, or continuous.

We begin by describing the method and showing how the Cox-based hazard ratio can be interpreted as an odds ratio. We present results of simulations to assess confidence interval coverage under the correct model and a null hypothesis of no association, comparing coverage percentages for our method with those for logistic regression and linear regression, based on both single impute substitution and multiple imputation. We then illustrate the reverse-scale Cox method by applying it to cross-sectional NHANES survey data to assess sex as a determinant of TCDD concentration in the US population. We also apply it to prospective cohort data examining the association between in utero PCB-28 exposure and psychomotor development, where predictor and outcome roles are reversed to accommodate nondetects. We conclude with some remarks about the strengths and limitations of the proposed method.

METHODS

Background

Survival methods are standard statistical tools typically used for analyzing censored positive-valued data such as event times. Most survival methods were originally developed to handle noninformatively right-censored outcomes, known only to exceed a given limit. For example, the death time for a patient in a prospective cohort study is right censored if the patient withdraws from the study or is still alive when the study concludes; in either situation, we simply know that the time to death exceeds the observed time on study.

Reversing the outcome scale to analyze left-censored data

Outcomes may also be left censored, known only to be less than some lower limit. Such data can be analyzed with methods designed for right-censored data by reversing the outcome scale (9). Specifically, and without loss of generality, one can choose a constant, say M, that equals (or exceeds) the largest observation, subtract all uncensored and left-censored outcomes from M, and treat those differences as uncensored and right-censored outcomes, respectively. Other than requiring that it equal or exceed all measured (or censored) values, the choice of M is arbitrary but will have no practical consequence on nonparametric or semiparametric inferences. This maneuver converts values that are left censored by the assay LOD to values that are right censored at $M - \text{LOD}$. Note that LODs can vary across participants, as occurs in our examples.

Following scale reversal, existing nonparametric methods can estimate the outcome distribution (or at least its tail) for a single group (9) or formally compare distributions for 2 or more groups (10). Here, we extend these ideas to use Cox regression (4) to adjust for confounders when assessing the association between a health measure and a quantitative exposure that is subject to substantial censoring.

Cox regression analysis of left-censored data

For a given individual, let T be the true concentration of an analyte, let Y be a health measure, and let $\mathbf{Z}$ be a vector of covariates. Define a censoring indicator, $\Delta = I_{(T > \text{LOD})}$, and a concentration variable, $C = \max(T, \text{LOD})$, where Δ is 1 if C equals the true concentration (i.e., T is large enough to be measured), and Δ is 0 if C equals the LOD (i.e., T is nondetectable). Suppose we have a sample of N observations, denoted $\{(c_i, \delta_i, y_i, \mathbf{z}_i)\colon i = 1, \ldots, N\}$, where $(c_i, \delta_i, y_i, \mathbf{z}_i)$ are the observed values of $(C, \Delta, Y, \mathbf{Z})$ for the ith individual $(i = 1, \ldots, N)$. Choose $M \geq \max(c_1, \ldots, c_N)$ to be a fixed constant that equals or exceeds the largest true concentration (and the largest LOD), and define $x_i = M - c_i$ $(i = 1, \ldots, N)$ to be the observed value of some hypothetical "reverse-scale" concentration represented by $X = M - C$.

Because most event-time software is designed for right-censored data, one can reverse the scale of left-censored data and apply standard software to the transformed data. Thus, one can use existing Cox regression software to analyze the data $\{(x_i, \delta_i, y_i, z_i): i = 1, \ldots, N\}$, treating each x_i as an outcome subject to right censoring, δ_i as a censoring indicator, and y_i and z_i as covariates ($i = 1, \ldots, N$). Because 1 of our examples involves NHANES data obtained from a multistage stratified cluster sample, we used the SURVEYPHREG procedure in SAS, version 9.3, software (SAS Institute, Inc., Cary, North Carolina), which incorporates information on sampling strata, clusters, and weights to provide appropriate standard errors when analyzing complex sample survey data. For our other example, we used the PHREG procedure in SAS, which is appropriate for nonsurvey Cox analyses.

Cox regression assumes that hazards are proportional to a baseline hazard; its parameters are log hazard ratios. A positive coefficient for Y implies that $M - T$ tends to be smaller for larger Y, and thus T tends to be larger. Let $F(t) = \Pr(T \leq t)$ be the cumulative distribution function for true concentration T, where t is a particular value of T, and let $f(t)$ be its density function. In contrast to the hazard function for T, given by $f(t) / [1 - F(t)]$, the hazard function for reverse-scale concentration, $M - T$, when rewritten in terms of T, is $f(t) / F(t)$.

If Y is binary, the hazard ratio parameter is interpretable as an odds ratio. When comparing people who are positive for the binary health measure ($Y = 1$) with those who are negative ($Y = 0$), the hazard ratio is $[f_1(t) / F_1(t)] / [f_0(t) / F_0(t)]$, where $f_y(t)$ and $F_y(t)$ are the respective density and distribution functions for persons with health measure $Y = y$ (for $y = 0,1$). Reversing the conditional probabilities, this hazard ratio can be rewritten as follows:

$$\begin{aligned}
&\frac{f_1(t)/F_1(t)}{f_0(t)/F_0(t)} \\
&\quad = \frac{\lim_{\varepsilon \to 0} Pr(t \leq T < t + \varepsilon \mid Y = 1)/Pr(T \leq t \mid Y = 1)}{\lim_{\varepsilon \to 0} Pr(t \leq T < t + \varepsilon \mid Y = 0)/Pr(T \leq t \mid Y = 0)} \\
&\quad = \frac{Pr(Y = 1 \mid T = t)/Pr(Y = 0 \mid T = t)}{Pr(Y = 1 \mid T \leq t)/Pr(Y = 0 \mid T \leq t)}, \qquad (1)
\end{aligned}$$

which is the odds of the health outcome at concentration t divided by the odds of the health outcome for the aggregate of concentrations at or below t. Thus, the usual proportional hazards assumption made in Cox regression corresponds to an assumption that this odds ratio is the same across all values of t (i.e., all concentrations). If Y is continuous, an analogous argument applies, leading to an odds ratio for a 1-unit change in Y.

SIMULATIONS

To assess the validity of our approach, we performed simulations to verify coverage for nominal 95% confidence intervals under a correctly specified null model. In addition to the Cox method, we evaluated several other methods across a range of scenarios. We simulated data with observations on a binary health outcome Y (e.g., case/control status), a quantitative confounder Z, and a concentration T subject to a high proportion of nondetects.

Design of simulations

In addition to our earlier notation, let R denote a reverse-scale version of T. There are n_1 cases ($Y = 1$), n_0 controls ($Y = 0$), and we set $N = n_0 + n_1$. We generated $\ln(Z)$ as normally distributed with mean μ_1 for cases ($Y = 1$), mean μ_0 for controls ($Y = 0$), and variance σ^2 for both. Simulating concentrations that have a proportional hazards relationship with the confounder and case/control status on the reverse concentration scale is tricky (see the Appendix for details). In general, the concentration distribution involves a hazard ratio and a baseline (i.e., covariate-free) distribution, where the former depends on Z and Y but not R, and the latter depends on R but not Z or Y. In our simulated data, the log hazard ratio for R was a linear function of Y and Z. We used a standard Weibull model, with scale parameter α and shape parameter γ, for the baseline distribution of R.

We wanted to assess performance under a null model of no association between exposure T (or equivalently R) and case/control status Y, so we set the coefficient of Y to 0. Without loss of generality, we set $\mu_0 = 0$ and regulated the dependence of Y on Z through μ_1, and we set $\alpha = \ln(2)$ to scale the baseline concentration distribution. We set $\beta = \ln(2)$ to obtain a hazard ratio of 2 for the effect of a 1-unit change in Z on R. We report results for lightly ($\mu_1 = 1$, $\sigma = 0.25$) and heavily ($\mu_1 = 0.5$, $\sigma = 1$) skewed confounder distributions, the former being nearly bell shaped, and for lightly ($\gamma = 5$) and heavily ($\gamma = 1$) skewed baseline distributions. We studied the following 5 pairs of sample sizes: $(n_0, n_1) = (100, 100)$, $(100, 400)$, $(400, 100)$, $(400, 400)$, and $(700, 700)$; and we imposed 3 fixed LODs to achieve average censoring proportions of 50%, 70%, and 90%.

We simulated 10,000 data sets for each combination of confounder distribution, baseline concentration distribution, and case/control sample sizes, and we calculated the percentage of those simulated studies for which the 95% confidence interval covered the true null value. For comparison, we also calculated coverages for 2 substitution methods: 1 based on logistic regression, with Y as the outcome and T and Z as covariates, and 1 based on linear regression, with T as the outcome and Y and Z as covariates. In both analyses, each nondetectable value of T was replaced by $\mathrm{LOD} / \sqrt{2}$ (11). We also calculated coverages for 2 multiple imputation methods based on the same logistic and linear models. Using the methods of Lubin et al. (12), we bootstrapped to estimate parameters under a lognormal model for T, and we generated 10 "fill-in" data sets (with imputed values for nondetects) for each of the 10,000 simulated data sets.

Simulation results

Table 1 reports confidence interval coverages for these 5 methods in each of the 20 null scenarios with a censoring proportion of 70%. The Cox method maintained the nominal 95% coverage in all scenarios. Linear regression often performed poorly, with coverage as low as 44% when using substitution for a lightly skewed (i.e., nearly symmetrical) covariate distribution and 56% when using multiple imputation

Table 1. Null Coverage Percentages for 95% Confidence Intervals by Shape of Covariate Distribution, Shape of Baseline Concentration Distribution, and Sample Sizes (With 70% Censoring)

Level of Skewness		Sample Size		Coverage Percentage for Regression Methods[a]				
Covariate Distribution	Concentration Distribution	n_0	n_1	Reverse Scale Cox	Linear with Substitution	Logistic with Substitution	Linear with MI	Logistic with MI
Light	Light	100	100	95	86	97	95	100
		100	400	95	70	96	95	100
		400	100	95	89	96	96	99
		400	400	95	64	95	95	99
		700	700	95	44	95	94	99
Light	Heavy	100	100	95	90	97	95	100
		100	400	95	80	96	94	100
		400	100	95	91	96	96	99
		400	400	95	75	95	95	99
		700	700	95	61	95	95	99
Heavy	Light	100	100	95	92	94	92	95
		100	400	95	91	94	88	93
		400	100	95	91	91	92	92
		400	400	95	84	93	79	90
		700	700	95	75	93	67	85
Heavy	Heavy	100	100	95	92	94	91	94
		100	400	95	87	94	84	92
		400	100	95	90	90	90	90
		400	400	95	77	91	73	86
		700	700	95	63	88	56	78

Abbreviations: LOD, limit of detection; MI, multiple imputation.

[a] Coverage percentages are based on 10,000 replicate data sets. The 5 methods are as follows: reverse-scale Cox regression, linear regression with substitution of LOD / $\sqrt{2}$ for nondetects, logistic regression with substitution of LOD / $\sqrt{2}$ for nondetects, linear regression with MI for nondetects, and logistic regression with MI for nondetects. See the Simulations section and the Appendix for details.

for a heavily skewed covariate distribution. Logistic regression fared better than linear regression but not as well as Cox regression, with coverages ranging from 88% to 97% when using substitution and from 78% to 100% when using multiple imputation. In some scenarios, for both linear regression and logistic regression, multiple imputation was worse than single impute substitution. Only the Cox method maintained the nominal 95% coverage in all 20 scenarios with 50% censoring (Web Table 1 available at http://aje.oxfordjournals.org/), and its coverages ranged from 94% to 96% when the censoring proportion rose to 90% (Web Table 2).

ILLUSTRATIVE EXAMPLES

Example 1: sex and TCDD exposure

Identifying determinants of environmental exposures is an important public health problem. Understanding what predicts exposure can lead to strategies to reduce exposure, and for causal models, provide causal background for selecting potential confounding variables. As 1 illustration of our method, we investigated sex as a possible determinant of TCDD exposure, using cross-sectional NHANES survey data from 1999 to 2004. Serum TCDD concentrations were determined by the Centers for Disease Control and Prevention, using high-resolution gas chromatography/isotope-dilution high-resolution mass spectrometry (13). Serum lipid concentrations were also determined, and total lipid concentration was estimated by the Akins summation method (14). We used TCDD concentrations expressed on a per-lipid basis (pg/g lipid), as supplied through NHANES. These data are anonymous and publicly available, and the National Institutes of Health Human Research Protection Program deemed this study exempt from further ethical or institutional review board review.

The original data set involved 7,433 participants who were probability sampled with oversampling of certain age, race, ethnicity, and income categories (15). Of these, the Centers for Disease Control and Prevention assessed TCDD for 5,002. We excluded pregnant women and any participant without information on sex or age, further reducing our sample size to 4,756. Within this subsample, 82% of the TCDD concentrations were below the LOD. The medians were 3.5 (range, 0.8–42.7) for the measured TCDD concentrations and 3.1 (range, 0.4–11.9) for the LODs.

We applied the proposed Cox method to these data, treating lipid-adjusted TCDD concentration as the (heavily censored) outcome and incorporating sex and age as covariates. For comparison, we also performed linear regression, with nondetects replaced by LOD / $\sqrt{2}$. Both analyses included a binary indicator for sex (1 = male, 0 = female); a simple quantitative term for age; and stratum, cluster, and weight variables to adjust for the clustered probability sampling (using SAS procedures SURVEYPHREG and SURVEYREG). We also performed analyses that instead used a categorical variable or a spline for age, and these analyses gave results very similar to our original analysis. We did not perform multiple imputation because of the lack of software able to impute "fill-in" data sets within the context of a complex survey sample.

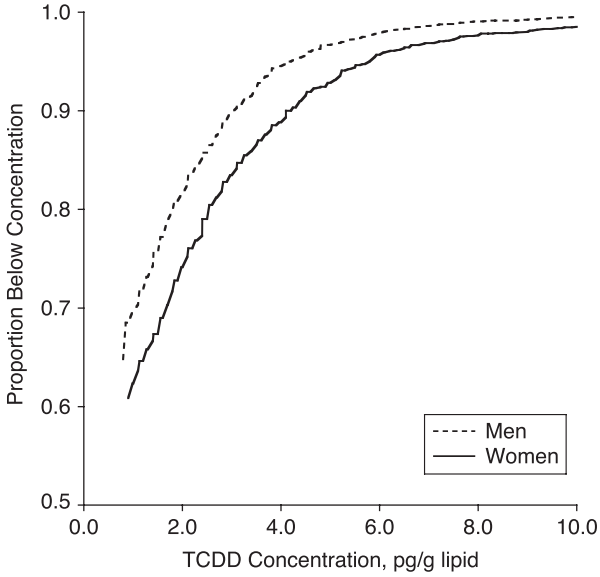

Figure 1. Differences in 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) concentration by sex, National Health and Nutrition Examination Survey, 1999–2004. Separate reverse Kaplan-Meier curves (9, 19) estimate the TCDD concentration distribution for men and women. At each (*x*,*y*) point along a curve, the *y*-value is the estimated sex-specific proportion of TCDD measurements below the TCDD concentration (in pg/g lipid) represented by the *x*-value. TCDD concentrations below the limit of detection were treated as left censored when constructing these curves, which are not adjusted for age or the survey design.

The Cox regression analysis showed both sex and age to be important determinants of TCDD exposure ($P < 0.0001$), with men having lower serum TCDD concentrations than women (Figure 1) and with TCDD concentrations increasing with age. On the reverse concentration scale, the estimated Cox regression coefficient for men was −0.45, and the estimated hazard ratio was $\exp(-0.45) = 0.64$ (95% confidence interval: 0.53, 0.76). In contrast, when linear regression was used with substitution of $\mathrm{LOD}/\sqrt{2}$ for each nondetect, age remained a significant predictor ($P < 0.0001$), but sex did not ($P = 0.30$).

Example 2: PCB exposure and neurodevelopment

To examine associations between in utero PCB exposures and neurodevelopment, Park et al. (8) used cohort data from mother-infant pairs residing in 2 districts of eastern Slovakia, with births between 2002 and 2004. In utero PCB exposure was estimated from concentrations measured in serum collected from the mother at the time of her child's birth. Fifteen PCB congeners were measured on a wet weight basis (as ng/mL) and then adjusted for serum lipids (as ng/g lipid) by the Akins method (14). Infant neurodevelopment was assessed using the Bayley Scales of Infant Development-II (16) at 16 months of age; our example focuses on results from the Psychomotor Development Index (PDI), 1 of 2 indices that comprise the Bayley Scales. To adjust for potential confounders when evaluating the association between individual maternal PCB concentrations and infant PDI scores, we included in our regression models district of residence (Michalovce or Svidnik), infant sex, maternal Raven score (a quantitative measure of nonverbal intelligence), and HOME score (a quantitative measure of quality and quantity of stimulation given to the infant at home). See the article by Park et al. (8) for details about this study, which was approved by the institutional review boards at the Slovak Medial University (Bratislava, Slovakia) and the University of California at Davis (Davis, California).

Of the 15 PCB congeners evaluated in maternal sera, Park et al. (8) included 6 in their linear regressions because those 6 "had most values above their LODs" and excluded the other 9 because of substantial censoring. Using 1 of the originally omitted congeners, PCB-28, we illustrate our Cox approach by assessing the association between maternal PCB-28 concentration and infant PDI score, adjusting for confounding. Data on PDI scores and the 4 potential confounders were available for 666 of the original 1,134 mother-infant pairs in the birth cohort, and among these, 59% of the maternal PCB-28 concentrations were below the LOD.

For comparison, we also performed 2 linear regressions, where nondetects either were replaced by $\mathrm{LOD}/\sqrt{2}$ or were multiply imputed with bootstrapping and 10 "fill-in" data sets, as described by Lubin et al. (12). Both linear regression analyses used PDI score as the outcome, with natural-log-transformed PCB-28 concentration as the exposure of interest and with the same 4 potential confounders. The Cox analysis strongly suggests that lower maternal PCB-28 concentrations are associated with higher infant PDI scores ($P = 0.01$), whereas the substitution ($P = 0.13$) and multiple imputation ($P = 0.12$) analyses provide only weak evidence for such an association. Infant PDI scores ranged from 51 to 148, and the estimated Cox hazard ratio was 0.88 (95% confidence interval: 0.79, 0.97) per 10-unit change. Thus, for 2 infants whose PDI scores differed by 10 points, the odds ratio was 0.88 for the mother of the higher-scoring child having a given PCB-28 concentration versus all lower concentrations. As for the linear regression analyses, the estimated changes in PDI score when doubling the PCB-28 concentration were −0.53 (95% confidence interval: −1.22, 0.15) with substitution and −0.47 (95% confidence interval: −1.06, 0.13) with multiple imputation. Note that the linear regression parameters have a different interpretation than the Cox regression parameter.

Though our approach is most useful with a high proportion of nondetects, we applied it to the 6 congeners investigated by Park et al. (8) to confirm that it also performs well with minimal censoring. The Cox method found significant associations between PDI score and the same 2 congeners that had been identified by Park et al. (8) using linear regression with substitution, and the signs of the regression coefficients in the Cox and linear models agreed.

DISCUSSION

When used to relate exposure biomarker analytes to a health outcome of interest, our method begins by reversing the concentration scale and then applies Cox regression with adjustment for potential confounders, that is, for factors

that may be causal "ancestors" of both the analyte concentration and the health outcome. Directed acyclic graph methods (5) can be used to identify those ancestors for inclusion in the model. Our simulations show that the approach is valid even with both confounding and extreme LOD censoring (i.e., a high proportion of nondetects), provided the model is correctly specified, whereas alternative approaches can perform poorly.

In our NHANES example, we were in fact testing whether sex influences TCDD concentrations, so sex was the cause and TCDD the outcome. In other applications, reversal of the roles of exposure and outcome may be needed, and in such settings the reverse-scale Cox method can be regarded as simply assessing association.

Our method has some important limitations. As with any complex model involving confounding covariates, misspecification of the model can invalidate inferences. The Cox analysis does not assume any parametric distributions, but it does require that the hazard functions for different covariate values be proportional across concentrations of the analyte under study. This assumption would be important to check, particularly because we have not here assessed the method's sensitivity to violations of that assumption. Other strategies for optimizing fit should be used, such as considering transformations for continuous confounders, though goodness of fit can be challenging to evaluate under the Cox model (17).

A second potential limitation is that characterizing a dose-response relationship may be difficult under the proposed approach, especially in situations in which the roles of the exposure and the outcome are reversed. Although estimation of a dose-response curve may be nearly impossible when LOD censoring is extreme, the proposed method will be useful for detecting risks from chemicals or metabolites of biomarkers under investigation. For substances identified, further testing could be done with a more sensitive assay to study dose-response relationships.

Other limitations relate to censoring. Our method may not be useful in a cohort study when the outcome of primary interest is the time to an event (e.g., death), and the predictor of interest is an exposure biomarker subject to a high proportion of nondetects. The potential problem with such a setting is that both the outcome and its predictor are subject to censoring. Also, Cox analysis typically assumes censoring is noninformative. In our application, censoring is inherently noninformative because knowing that the analyte concentration falls below a particular LOD should not tell us more than that about its actual value.

The proposed method has some notable strengths. Unlike approaches that discard nondetects or analyze detect/nondetect dichotomies, our method allows full use of the available data for a quantitative exposure biomarker. Unlike approaches that substitute specific values (e.g., LOD / 2 or LOD / $\sqrt{2}$) for nondetects, our method does not assume that unknown values are known and thus should not experience the same biases (both in estimation and testing) and underestimations of variability.

Application to matched case-control studies is also possible with the proposed method. If frequency matching has been applied, the usual adjustments for the matching factors must be included in the model. The adjusted odds ratio from the Cox model can then be viewed as equivalent to one from a prospective formulation, as a consequence of results by Prentice and Pyke (18). On the other hand, if pair matching has been performed on the basis of nonquantified factors (e.g., with each case matched to a friend control or a sibling control), then an analyte with a small number of known concentrations cannot be readily studied using our method. However, such a design presents an intractable and possibly insurmountable challenge for other approaches, too. For example, suppose 80% are nondetects. Even if analyte concentrations are independent within pairs, 64% of pairs would be completely noninformative, and only 4% of pairs would have 2 measured concentrations to compare. Consequently, we caution against such a design if the number of known analyte concentrations is small.

In summary, a reverse-scale Cox model can be used to assess a confounder-adjusted association between a health outcome (or an exposure determinant) and an exposure biomarker whose concentrations often fall below its assay's LOD. The method fully uses quantifiable analyte measures, appropriately treats nondetects as censored, avoids specific parametric assumptions about the concentration distribution, and enables adjustment for confounders.

ACKNOWLEDGMENTS

Author affiliations: Biostatistics Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina (Gregg E. Dinse, Clarice R. Weinberg); Department of Public Health Sciences, Division of Epidemiology, University of Rochester School of Medicine and Dentistry, Rochester, New York (Todd A. Jusko); Health Sciences Research Group, Survey Research Associates International, Durham, North Carolina (Lindsey A. Ho); Department of Biology, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania (Kaushik Annam); Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania (Kaushik Annam); Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland (Barry I. Graubard); Department of Public Health Sciences, School of Medicine, University of California Davis, Davis, California (Irva Hertz-Picciotto); Environmental Autoimmunity Group, Program of Clinical Research, National Institute of Environmental Health Sciences, National Institutes of Health Clinical Research Center, Bethesda, Maryland (Frederick W. Miller); and Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, Michigan (Brenda W. Gillespie).

This research was supported in part by the intramural research program of the National Institutes of Health, National Institute of Environmental Health Sciences (projects ZIA ES040006 and ZIA ES101074) and by the National Institutes of Health (grants K12 ES019852, P30 ES001247, and R01 CA096525).

We are grateful to Drs. G. Kissling, M. Longnecker, and R. Morris for helpful discussions, to C. Co and R. Jaramillo for programming assistance, and to members of the National

Health and Nutrition Examination Study Autoimmunity Study Group (Drs. M. Satoh, E. Chan, K. Rose, C. Parks, R. Cohn, N. Walker, D. Germolec, I. Whitt, L. Birnbaum, and D. Zeldin) for initiating the studies that motivated much of this research.

Conflict of interest: none declared.

REFERENCES

1. Browne RW, Whitcomb BW. Procedures for determination of detection limits: application to high-performance liquid chromatography analysis of fat-soluble vitamins in human serum. *Epidemiology*. 2010;21(suppl 4):S4–S9.
2. Nie L, Chu H, Liu C, et al. Linear regression with an independent variable subject to a detection limit. *Epidemiology*. 2010;21(suppl 4):S17–S24.
3. Arunajadai SG, Rauh VA. Handling covariates subject to limits of detection in regression. *Environ Ecol Stat*. 2012;19(3):369–391.
4. Cox DR. Regression models and life-tables. *J Roy Statist Soc Ser B*. 1972;34(2):187–220.
5. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10(1):37–48.
6. Satoh M, Chan EK, Ho LA, et al. Prevalence and sociodemographic correlates of antinuclear antibodies in the United States. *Arthritis Rheum*. 2012;64(7):2319–2327.
7. Gallagher CM, McElroy AE, Smith DM, et al. Polychlorinated biphenyls, mercury, and antinuclear antibody positivity, NHANES 2003–2004. *Int J Hyg Environ Health*. 2013;216(6):721–727.
8. Park HY, Hertz-Picciotto I, Sovcikova E, et al. Neurodevelopmental toxicity of prenatal polychlorinated biphenyls (PCBs) by chemical structure and activity: a birth cohort study. *Environ Health*. 2010;9:51.
9. Gillespie BW, Chen Q, Reichert H, et al. Estimating population distributions when some data are below a limit of detection by using a reverse Kaplan-Meier estimator. *Epidemiology*. 2010;21(suppl 4):S64–S70.
10. Zhang D, Fan C, Zhang J, et al. Nonparametric methods for measurements below detection limit. *Stat Med*. 2009;28(4):700–715.
11. Hornung RW, Reed LD. Estimation of average concentration in the presence of nondetectable values. *Appl Occup Environ Hyg*. 1990;5(1):46–51.
12. Lubin JH, Colt JS, Camann D, et al. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environ Health Perspect*. 2004;112(17):1691–1696.
13. Patterson DG, Hampton L, Lapeza CR, et al. High-resolution gas chromatographic/high-resolution mass spectrometric analysis of human serum on a whole-weight and lipid basis for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. *Anal Chem*. 1987;59(15):2000–2005.
14. Akins JR, Waldrep K, Bernert JT. The estimation of total serum lipids by a completely enzymatic 'summation' method. *Clin Chim Acta*. 1989;184(3):219–226.
15. Mohadjer L, Curtin LR. Balancing sample design goals for the National Health and Nutrition Examination Survey. *Surv Methodol*. 2008;34(1):119–126.
16. Bayley N. *Bayley Scales of Infant Development*. San Antonio, TX: The Psychological Corporation; 1993.
17. Hosmer DW, Lemeshow S, May S. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2008.
18. Prentice RL, Pyke R. Logistic disease incidence models and case-control studies. *Biometrika*. 1979;66(3):403–411.
19. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):457–481.

APPENDIX

Define a binary outcome Y, a quantitative confounder Z, and a quantitative reverse-scale concentration R. We simulated various null data under a proportional hazards model for R, as assumed by the Cox method. True concentration T, observed concentration $C = \max(T, \text{LOD})$, and censoring indicator $\Delta = I_{(T > \text{LOD})}$ can be obtained from R and the LOD.

The usual formulation of the proportional hazards model (4, 17) implies

$$\ln[S(r \mid y, z)] = \exp(\phi y + \beta z) \cdot \ln[S(r \mid 0, 0)], \qquad \text{(A1)}$$

where $S(r \mid y,z) = Pr(R > r \mid Y = y, Z = z)$ is the "survival" (or upper-tail distribution) function for R given Y and Z, and $S(r \mid 0,0)$ is a baseline survival function that depends only on R. Regression coefficients ϕ and β are log hazard ratios that gauge the effect on R of Y and Z, respectively. We used the following Weibull baseline survival function:

$$S(r \mid 0, 0) = \exp(-\alpha r^{\gamma}), \qquad \text{(A2)}$$

where α is a scale parameter, and γ is a shape parameter.

We simulated "observed" data $\{(y_i, z_i, c_i, \delta_i): i = 1, \ldots, N\}$ for N persons indexed by i as follows. We assigned $y_i = 0$ for n_0 persons and $y_i = 1$ for n_1 persons, where $N = n_0 + n_1$. Given y_i, we sampled $\ln(z_i)$ from a normal distribution with mean μ_0 if $y_i = 0$, mean μ_1 if $y_i = 1$, and variance σ^2 for either value of y_i and then exponentiated to get z_i. We set $\mu_0 = 0$ and controlled the dependence of Z on Y through μ_1; we chose pairs ($\mu_1 = 1, \sigma = 0.25$) and ($\mu_1 = 0.5, \sigma = 1$) to produce lightly and heavily skewed distributions for Z, respectively. Given y_i and z_i, we generated r_i using formulas A1 and A2. We set $\phi = 0$ to enforce the null hypothesis of no association between Y and R (or equivalently T) and $\beta = \ln(2)$ to obtain a hazard ratio of 2 for the effect of a 1-unit change in Z on R. With respect to the baseline survival function, we set $\alpha = \ln(2)$ to yield a median of 1, and we chose shapes $\gamma = 5$ and $\gamma = 1$ to produce light and heavy skewness, respectively. We created a reverse-scale limit of detection U by determining the value that yields an average $S(U \mid y_i,z_i)$, in large samples, equal to the desired censoring rate. Then we defined $V = \max(r_1, \ldots, r_N)$ and $\text{LOD} = V - U$. Finally, we set ($\delta_i = 1$, $c_i = V - r_i$) if $r_i < U$, and we set ($\delta_i = 0$, $c_i = \text{LOD}$) otherwise. This procedure produced "observed" data of the form $\{(y_i, z_i, c_i, \delta_i): i = 1, \ldots, N\}$ for the Cox analysis. The linear and logistic regression analyses, based on substitution or multiple imputation, required data on T. If $\delta_i = 1$, then T was uncensored, and we set $t_i = c_i$ in either case. If $\delta_i = 0$, the substitution approach set $t_i = \text{LOD} / \sqrt{2}$, and the imputation approach used missing data methods to impute a value of t_i from the estimated distributions of the observed data (12).