Assessing the Sensitivity of Regression Results to Unmeasured Confounders in Observational Studies

D. Y. Lin,[1,*] B. M. Psaty,[2] and R. A. Kronmal[1]

[1]Department of Biostatistics, University of Washington,
Box 357232, Seattle, Washington 98195, U.S.A.
[2]Departments of Medicine and Epidemiology, University of Washington,
Seattle, Washington 98195, U.S.A.

SUMMARY

This paper presents a general approach for assessing the sensitivity of the point and interval estimates of the primary exposure effect in an observational study to the residual confounding effects of unmeasured variables after adjusting for measured covariates. The proposed method assumes that the true exposure effect can be represented in a regression model that includes the exposure indicator as well as the measured and unmeasured confounders. One can use the corresponding reduced model that omits the unmeasured confounder to make statistical inferences about the true exposure effect by specifying the distributions of the unmeasured confounder in the exposed and unexposed groups along with the effects of the unmeasured confounder on the outcome variable. Under certain conditions, there exists a simple algebraic relationship between the true exposure effect in the full model and the apparent exposure effect in the reduced model. One can then estimate the true exposure effect by making a simple adjustment to the point and interval estimates of the apparent exposure effect obtained from standard software or published reports. The proposed method handles both binary response and censored survival time data, accommodates any study design, and allows the unmeasured confounder to be discrete or normally distributed. We describe applications to two major medical studies.

1. Introduction

Well-conducted randomized controlled experiments have been widely recognized as the decisive method for assessing exposures, treatments or policies. Randomized experiments are not feasible in many circumstances because of ethical, economic, and other constraints. Instead, observational studies, including cohort and case–control studies, are commonly undertaken. Because subjects are not assigned to exposure groups at random in observational studies, the apparent association between exposure and outcome may result from some unknown or unmeasured covariates that are associated with both exposure and outcome. Thus, it would be desirable to investigate how the findings of an observational study may be affected by certain variations in the assumptions about the unmeasured confounders. If the conclusions are insensitive over a wide range of plausible assumptions, then the number of interpretations of the observational data is reduced, and the causal conclusions become more defensible.

This idea of sensitivity analysis can be traced back to Cornfield et al. (1959), who responded to Fisher (1958)'s constitutional hypothesis that the apparent association between cigarette smoking and lung cancer could be explained by the confounding effects of a genotype that predisposed to both smoking and lung cancer. A centerpiece of their argument was a theorem due to Cornfield: for a confounder to explain a relative risk of a given magnitude, say, r, this confounder has to be

r times more prevalent among the exposed than among the unexposed. Because the lung cancer relative risk was approximately 10 for cigarette smokers versus nonsmokers and increased to 20-fold for heavy smokers, the existence of such a confounder seemed implausible. This line of statistical reasoning played a critical role in the 1964 Surgeon General's Report concluding that "Cigarette smoking is causally related to lung cancer in men."

Bross (1966) and Schlesselman (1978) obtained results similar to Cornfield's inequality. Rosenbaum (1995, Chapter 4) provided the bounds for the significance levels of several common two-sample tests and for Hodges–Lehmann estimates. In addition, Rosenbaum and Rubin (1983a) discussed the sensitivity analysis on a point estimate for the difference between two success probabilities in a $2 \times 2 \times S$ contingency table. The method of Rosenbaum and Rubin (1983a) specifies stratum-specific sensitivity parameters, consisting of the overall prevalence of the binary confounder, the confounder's effect on the exposure and the confounder's effect on the binary outcome. The method then provides a range of point estimates corresponding to various combinations of the sensitivity parameters. By contrast, the methods of Cornfield et al. (1959) and Rosenbaum (1995) specify only a single sensitivity parameter and give the corresponding bound for the exposure effect that caters to the least favorable or most conservative values of the remaining parameters.

In this paper, we develop a general yet simple method of sensitivity analysis that shares the spirit of the published methods, especially that of Rosenbaum and Rubin (1983a). Formulating the problem more broadly than Rosenbaum and Rubin, we provide an alternative solution that applies to a variety of statistical models and data structures. Specifically, we assume that the exposure effect and the effects of measured and unmeasured confounders can be formulated through a regression model. This formulation implies a reduced model that excludes the unmeasured confounder. It is possible to estimate the true exposure effect in the full model by fitting the reduced model to the observable data provided that one specifies the distribution of the unmeasured confounder in each exposure group as well as the effects of the unmeasured confounder on the outcome variable. As a matter of fact, given these two sets of sensitivity parameters, there exists a simple algebraic relationship between the true exposure effect in the full model and the apparent exposure effect in the corresponding reduced model if the unmeasured and measured confounders are approximately independent within each exposure group. In such situations, we can draw inferences about the true exposure effect by making simple adjustments to the point and interval estimates of the apparent exposure effect, which are obtained either from the standard software package or from the published paper.

The proposed method offers several advantages over the existing ones. First, it applies to both binary response and censored survival time data and to the corresponding logistic and proportional hazards models. Survival time or failure time is an important outcome variable in many observational cohort studies, but this endpoint has not been carefully studied in the sensitivity analysis literature, although Rosenbaum's method of permutation inferences may be applied to log-rank and Gehan–Wilcoxon tests. Second, the proposed method allows adjustment for any type of measured confounders, discrete or continuous, whereas the published methods can at most accommodate stratification on categorical variables. Third, the unmeasured confounders may be binary or normally distributed; the existing literature is confined to a bounded confounder. Finally, we provide simple and explicit formulas to adjust both the point and the interval estimates of the apparent exposure effect in the reduced model. These formulas hold for both binary and survival endpoints and are invariant across study designs. For binary endpoint with a binary unmeasured confounder and with no adjustment of measured covariates, our results reduce to those of Cornfield, Schlesselman, and others.

2. Binary Outcome

Let Y be a binary response variable, taking the value 1 if the event of interest (e.g., disease or death) occurs, and let X be the exposure indicator, taking the value 1 for the exposed. Also, let $\mathbf{Z}$ be a set of measured covariates, and let U be an unmeasured confounder. We first assume that the response probability is related to X, $\mathbf{Z}$, and U through the log-linear model

$$\Pr(Y = 1 \mid X, \mathbf{Z}, U) = \exp\left(\alpha + \beta X + \gamma_X U + \boldsymbol{\theta}' \mathbf{Z}\right), \qquad (2.1)$$

where α, β, γ_X $(X = 0, 1)$, and $\boldsymbol{\theta}$ are unknown regression parameters. Note that γ_0 and γ_1 pertain to the effects of U for the unexposed and exposed, respectively. The parameterization $\beta X + \gamma_X U$ is the same as $\beta X + \gamma_0 U + (\gamma_1 - \gamma_0) X U$. If $\gamma_0 \neq \gamma_1$, then β is the main effect of X under a model in which X interacts with U so that β cannot be interpreted in isolation. It suffices most practical purposes to set $\gamma_0 = \gamma_1$.

Since U is unmeasured, one is forced to fit the reduced model

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \exp\left(\alpha^* + \beta^* X + \boldsymbol{\theta}^{*\prime} \mathbf{Z}\right), \tag{2.2}$$

where α^*, β^*, and $\boldsymbol{\theta}^*$ are potentially different from α, β, and $\boldsymbol{\theta}$ of model (2.1). We refer to β and β^* as the true and apparent exposure effects, respectively. Since β^* can be directly estimated from the observable data whereas β cannot, it is of great interest to ascertain the relationship between β^* and β.

Model (2.1) is slightly different from the logistic regression model,

$$\Pr(Y = 1 \mid X, \mathbf{Z}, U) = \frac{\exp\left(\alpha + \beta X + \gamma_X U + \boldsymbol{\theta}^{\prime} \mathbf{Z}\right)}{1 + \exp\left(\alpha + \beta X + \gamma_X U + \boldsymbol{\theta}^{\prime} \mathbf{Z}\right)}; \tag{2.3}$$

i.e.,

$$\mathrm{logit}\{\Pr(Y = 1 \mid X, \mathbf{Z}, U)\} = \alpha + \beta X + \gamma_X U + \boldsymbol{\theta}^{\prime} \mathbf{Z}.$$

For rare events, the right side of (2.1) is a good approximation to that of (2.3). If $\mathbf{Z}$ is excluded from models (2.1) and (2.3) and U is binary, then these two models are saturated and do not impose any restrictions on the effects of X and U on Y. In this case, models (2.1) and (2.3) are equivalent. In the presence of $\mathbf{Z}$, model (2.1) is mathematically more tractable than (2.3). Thus, it is instructive to consider model (2.1) first.

2.1 *Log-Linear Regression*

Let $F(u \mid X, \mathbf{Z})$ be the distribution function of U given X and $\mathbf{Z}$. By the law of conditional expectation,

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \int_{-\infty}^{\infty} \Pr(Y = 1 \mid X, \mathbf{Z}, u) dF(u \mid X, \mathbf{Z}), \tag{2.4}$$

which becomes

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \exp\left(\alpha + \beta X + \boldsymbol{\theta}^{\prime} \mathbf{Z}\right) \int_{-\infty}^{\infty} e^{\gamma_X u} dF(u \mid X, \mathbf{Z}) \tag{2.5}$$

under model (2.1). We will evaluate the right side of the above equation for the cases of binary and normally distributed U.

Binary confounder. Suppose that U is binary such that $F(u \mid X, \mathbf{Z})$ is a Bernoulli distribution with success probability $P_{X,\mathbf{Z}}$. Then equation (2.5) becomes

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \exp\left(\alpha + \beta X + \boldsymbol{\theta}^{\prime} \mathbf{Z}\right) \left\{ e^{\gamma_X} P_{X,\mathbf{Z}} + (1 - P_{X,\mathbf{Z}}) \right\}. \tag{2.6}$$

For simplicity, we assume that U is independent of $\mathbf{Z}$ conditional on X, which implies that $P_{X,\mathbf{Z}}$ does not depend on $\mathbf{Z}$. Under $P_{X,\mathbf{Z}} = P_X$, equation (2.6) may be written as

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \exp\left[\alpha + \log\left\{e^{\gamma_0} P_0 + (1 - P_0)\right\} + \left\{\beta + \log \frac{e^{\gamma_1} P_1 + (1 - P_1)}{e^{\gamma_0} P_0 + (1 - P_0)}\right\} X + \boldsymbol{\theta}^{\prime} \mathbf{Z}\right]. \tag{2.7}$$

Comparing the right sides of (2.2) and (2.7), we find that

$$\beta = \beta^* - \log \frac{e^{\gamma_1} P_1 + (1 - P_1)}{e^{\gamma_0} P_0 + (1 - P_0)} \tag{2.8}$$

or

$$R = R^*/A, \tag{2.9}$$

where $R = e^{\beta}$, $R^* = e^{\beta^*}$, and

$$A = \frac{\Gamma_1 P_1 + (1 - P_1)}{\Gamma_0 P_0 + (1 - P_0)}$$

with $\Gamma_1 = e^{\gamma_1}$ and $\Gamma_0 = e^{\gamma_0}$. Note that R and R^* are, respectively, the true and apparent relative risks of disease associated with X, while Γ_1 and Γ_0 are the relative risks of disease associated with U among the exposed and unexposed, respectively. Formula (2.9) shows succinctly how the prevalences of the unmeasured confounder, P_1 and P_0, and its relative risks of disease, Γ_1 and Γ_0, among the exposed and unexposed determine the relationship between the true and apparent relative risks of disease associated with the exposure.

It is of most interest to examine the adjustment factor A when $P_1 > P_0$ and $\Gamma_1 \geq \Gamma_0 > 1$. In such a setting, A is bounded above by Γ_1; this bound is achieved when $P_0 = 0$ and $P_1 = 1$. Under $\Gamma_0 = \Gamma_1 = \Gamma$, A is also bounded above by P_1/P_0; it approaches this bound as $\Gamma \to \infty$. In applications, it is more informative to evaluate A for specific values of P_0, P_1, Γ_0, and Γ_1 rather than using the upper bounds.

Normal confounder. Suppose that, conditional on X and $\mathbf{Z}$, the confounder U is normally distributed with mean $\mu_{X,\mathbf{Z}}$ and unit variance. The standardization of U entails that γ_0 and γ_1 are standardized regression coefficients. A simple integral calculation yields

$$\int_{-\infty}^{\infty} e^{\gamma_X u} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(u - \mu_{X,\mathbf{Z}})^2}{2}\right\} du = \exp\left(\gamma_X \mu_{X,\mathbf{Z}} + 0.5\gamma_X^2\right),$$

which entails that equation (2.5) is

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \exp\left(\alpha + \beta X + \boldsymbol{\theta}'\mathbf{Z} + \gamma_X \mu_{X,\mathbf{Z}} + 0.5\gamma_X^2\right). \tag{2.10}$$

As with the binary confounder, considerable simplification arises when U is conditionally independent of $\mathbf{Z}$ given X. Under $\mu_{X,\mathbf{Z}} = \mu_X$, equation (2.10) can be written as

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \exp\left[\alpha + \gamma_0 \mu_0 + 0.5\gamma_0^2 + \left\{\beta + (\gamma_1 \mu_1 - \gamma_0 \mu_0) + 0.5(\gamma_1^2 - \gamma_0^2)\right\} X + \boldsymbol{\theta}'\mathbf{Z}\right].$$

Thus,

$$\beta = \beta^* - \left\{(\gamma_1 \mu_1 - \gamma_0 \mu_0) + 0.5\left(\gamma_1^2 - \gamma_0^2\right)\right\}, \tag{2.11}$$

or

$$R = R^* / \exp\left\{(\gamma_1 \mu_1 - \gamma_0 \mu_0) + 0.5\left(\gamma_1^2 - \gamma_0^2\right)\right\}. \tag{2.12}$$

Further simplification arises when $\gamma_0 = \gamma_1 = \gamma$, in which case (2.11) and (2.12) reduce to

$$\beta = \beta^* - \gamma\delta, \tag{2.13}$$

$$R = R^* / \Gamma^\delta, \tag{2.14}$$

where $\delta = \mu_1 - \mu_0$. Under $\gamma_0 = \gamma_1$, equations (2.13)–(2.14) hold even if U is not conditionally independent of $\mathbf{Z}$, but the effects of X and $\mathbf{Z}$ on $\mu_{X,\mathbf{Z}}$ must be additive; i.e., $\mu_{X,\mathbf{Z}} = \mu_X + q(\mathbf{Z})$, where q is some arbitrary function of $\mathbf{Z}$. Note that the adjustment factor Γ^δ in (2.14) depends on the difference between μ_1 and μ_0, not on their actual values.

Formulas (2.8) and (2.9) and (2.11) to (2.14) enable one to conduct the sensitivity analysis in an extremely simple fashion. The sensitivity parameters consist of $(P_0, P_1, \Gamma_0, \Gamma_1)$ and $(\mu_0, \mu_1, \Gamma_0, \Gamma_1)$ in the binary and normal cases, respectively. Since β^* and consequently R^* can be directly estimated by fitting model (2.2) to the observed data on $(Y, X, \mathbf{Z})$, we can use formulas (2.8) and (2.9) and (2.11) to (2.14) to estimate β and R after specifying the sensitivity parameters. Denote the estimators of β^*, β, R^*, and R by $\hat{\beta}^*$, $\hat{\beta}$, $\hat{R}^*$, and $\hat{R}$. Equations (2.8) and (2.11) show that, when the sensitivity parameters are fixed at specific values, the variance of $\hat{\beta}$ is the same as that of $\hat{\beta}^*$. The latter can be estimated by the standard estimation theory. The fact that, for fixed values of the sensitivity parameters, the estimators $\hat{\beta}$ and $\hat{\beta}^*$ have the same variance implies that the adjustment factors for the limits of the confidence intervals for β are the same as that of the point estimate. In other words, formulas (2.8) and (2.9) and (2.11) to (2.14) apply to both the point and the interval estimates of β. Furthermore, all these results hold for any study design: prospective or retrospective, matched or unmatched.

If the conditional independence of U and $\mathbf{Z}$ given X is not true, then model (2.2) generally cannot hold, and there is no explicit relationship between β and β^*. Nevertheless, when the sensitivity parameters $(\gamma_X, P_{X,\mathbf{Z}})$ and $(\gamma_X, \mu_{X,\mathbf{Z}})$ are specified in the binary and normal cases, respectively, the right sides of models (2.6) and (2.10) involve only the measured variables X and $\mathbf{Z}$ and the corresponding regression parameters α, β and $\boldsymbol{\theta}$. One may then obtain $\hat{\beta}$ and its variance estimator by applying the maximum likelihood method or other requisite methods to models (2.6) and (2.10). In fact, given any specification of $F(u \mid X, \mathbf{Z})$ and γ_X, one can make inferences about β by fitting model (2.5) directly to the observable data $(Y_i, X_i, \mathbf{Z}_i)$ $(i = 1, \ldots, n)$.

2.2 Logistic Regression

Under the logistic model given in (2.3), equation (2.4) becomes

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \exp\left(\alpha + \beta X + \boldsymbol{\theta}'\mathbf{Z}\right) \int_{-\infty}^{\infty} \frac{e^{\gamma_X u}}{1 + \exp\left(\alpha + \beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}\right)} dF(u \mid X, \mathbf{Z}).$$

(2.15)

In the presence of covariates $\mathbf{Z}$, equation (2.15) generally does not reduce to a logistic model even if $F(u \mid X, \mathbf{Z})$ has a simple form. In other words, the following model usually fails:

$$\Pr(Y = 1 \mid X, \mathbf{Z}) = \frac{\exp(\alpha^* + \beta^* X + \boldsymbol{\theta}^{*'}\mathbf{Z})}{1 + \exp\left(\alpha^* + \beta^* X + \boldsymbol{\theta}^{*'}\mathbf{Z}\right)};$$

i.e.,

$$\mathrm{logit}\{\Pr(Y = 1 \mid X, \mathbf{Z})\} = \alpha^* + \beta^* X + \boldsymbol{\theta}^{*'}\mathbf{Z}. \tag{2.16}$$

Nonetheless, once γ_X and $F(u \mid X, \mathbf{Z})$ are specified, the right side of equation (2.15) involves only the measured variables X and $\mathbf{Z}$ and the corresponding regression parameters α, β, and $\boldsymbol{\theta}$. Thus, the likelihood type method can be applied to model (2.15) to make inferences about β, as will be elaborated in Section 5.

It is of practical importance to ascertain under what circumstances model (2.16) provides a reasonable approximation to model (2.15) and whether simple relationships such as (2.8) and (2.11) exist for the logistic regression. Straightforward algebraic manipulations show that model (2.15) is

$$\mathrm{logit}\{\Pr(Y = 1 \mid X, \mathbf{Z})\} = \alpha + \beta X + \boldsymbol{\theta}'\mathbf{Z} + g(X, \mathbf{Z}), \tag{2.17}$$

where

$$g(X, \mathbf{Z}) = \log \frac{e^{\gamma_X}\{1 + \exp(\alpha + \beta X + \boldsymbol{\theta}'\mathbf{Z})\}P_{X,\mathbf{Z}} + \{1 + \exp(\alpha + \beta X + \gamma_X + \boldsymbol{\theta}'\mathbf{Z})\}(1 - P_{X,\mathbf{Z}})}{\{1 + \exp(\alpha + \beta X + \boldsymbol{\theta}'\mathbf{Z})\}P_{X,\mathbf{Z}} + \{1 + \exp(\alpha + \beta X + \gamma_X + \boldsymbol{\theta}'\mathbf{Z})\}(1 - P_{X,\mathbf{Z}})}$$

for the binary confounder, and

$$g(X, \mathbf{Z}) = \gamma_X \mu_{X,\mathbf{Z}} + 0.5\gamma_X^2$$
$$+ \log \frac{\int_{-\infty}^{\infty} \left\{1 + \exp\left(\alpha + \gamma_X^2 + \beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}\right)\right\}^{-1} \exp\left\{-\frac{(u - \mu_{X,\mathbf{Z}})^2}{2}\right\} du}{\int_{-\infty}^{\infty} \{1 + \exp(\alpha + \beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z})\}^{-1} \exp\left\{-\frac{(u - \mu_{X,\mathbf{Z}})^2}{2}\right\} du}$$

for the normal confounder. If the event is rare or if $|\gamma_X|$ is small, $g(X, \mathbf{Z}) \approx \log\{e^{\gamma_X} P_{X,\mathbf{Z}} + (1 - P_{X,\mathbf{Z}})\}$ in the binary case and $g(X, \mathbf{Z}) \approx \gamma_X \mu_{X,\mathbf{Z}} + 0.5\gamma_X^2$ in the normal case. It then follows that, under $P_{X,\mathbf{Z}} = P_X$ and $\mu_{X,\mathbf{Z}} = \mu_X$, equation (2.17) becomes

$$\mathrm{logit}\{\Pr(Y = 1 \mid X, \mathbf{Z})\} \approx \left[\alpha + \log\left\{e^{\gamma_0} P_0 + (1 - P_0)\right\}\right]$$
$$+ \left\{\beta + \log \frac{e^{\gamma_1} P_1 + (1 - P_1)}{e^{\gamma_0} P_0 + (1 - P_0)}\right\} X + \boldsymbol{\theta}'\mathbf{Z} \tag{2.18}$$

and

$$\mathrm{logit}\{\Pr(Y = 1 \mid X, \mathbf{Z})\} \approx \left(\alpha + \gamma_0 \mu_0 + 0.5\gamma_0^2\right)$$
$$+ \left\{\beta + (\gamma_1 \mu_1 - \gamma_0 \mu_0) + 0.5\left(\gamma_1^2 - \gamma_0^2\right)\right\} X + \boldsymbol{\theta}'\mathbf{Z} \tag{2.19}$$

in the binary and normal cases, respectively. Models (2.18) and (2.19) take the form of (2.16) and imply that formulas (2.8) and (2.9) and (2.11) to (2.14) hold approximately for the logistic regression. Note that equations (2.13) and (2.14) hold approximately as long as the effects of X and $\mathbf{Z}$ on $\mu_{X,\mathbf{Z}}$ are additive. The conditional independence of U and $\mathbf{Z}$ given X required by formulas (2.8), (2.9), (2.11), and (2.12) will be discussed in Section 5.

When model (2.3) and consequently model (2.15) hold and model (2.16) does not, the maximum likelihood estimator $\hat{\beta}^*$ under model (2.16) converges to some well-defined limit, say, $\tilde{\beta}^*$, which minimizes the Kullback–Leibler distance between the true and assumed models (White, 1982). It is worthwhile to assess how close $\tilde{\beta}^*$ is to the value of β^* determined by equation (2.8), (2.11), or (2.13), that is, whether the substitution of $\hat{\beta}^*$ under model (2.16) for β^* in formula (2.8), (2.11), or (2.13) would lead to an approximately unbiased estimator of β. Unfortunately, no explicit analytic expression for $\tilde{\beta}^*$ exists, so we resort to Monte Carlo simulation, as described below.

In our simulation studies, the binary responses were generated from model (2.3) with $\beta = \theta = 1$, $\gamma_0 = \gamma_1$, X taking the value 1 for 20% of the subjects and Z being normal with mean X and variance 1. The intercept term α was set at different values to yield certain overall event rates. There were 1000 simulation samples, each consisting of 1000 subjects. For each simulation sample, we calculated the maximum likelihood estimate $\hat{\beta}^*$ under model (2.16) and replaced β^* in formula (2.8) or (2.13) with $\hat{\beta}^*$ to obtain $\hat{\beta}$ corresponding to specific values of the sensitivity parameters. In the case of binary U, we first set $P_0 = 0.2$ and $P_1 = 0.9$. Under $\Gamma = (2, 3, 4)$, the Monte Carlo estimates for the sampling means of $\hat{\beta}$ were approximately $(1.01, 1.02, 1.03)$ when the overall event rate was 10% and approximately $(1.02, 1.05, 1.08)$ when the overall event rate was 50%. For $P_0 = 0.4$ and $P_1 = 0.6$, the sampling means of $\hat{\beta}$ under $\Gamma = (2, 3, 4)$ were estimated at $(1.00, 0.98, 0.96)$ and $(0.98, 0.96, 0.94)$ when the overall event rates were 10% and 50%, respectively. Clearly, all these sampling means are very close to 1, the true value of β. In the case of normal U, we set the overall event rate at 10%. When $\mu_0 = 0$ and $\mu_1 = 1$, the Monte Carlo estimates for the sampling means of $\hat{\beta}$ were approximately $(0.99, 0.96, 0.90, 0.80)$ under $\gamma = (0.25, 0.5, 0.75, 1)$. The corresponding estimates were approximately $(0.99, 0.94, 0.86, 0.75)$ when $\mu_0 = 0$ and $\mu_1 = 2$. The bias of $\hat{\beta}$ became smaller as the event rate decreased.

These numerical results suggest that, for binary U, one may obtain $\hat{\beta}^*$ and its variance estimate from an existing logistic regression software and then use formulas (2.8) and (2.9) to perform the sensitivity analysis; for normal U, formulas (2.11) to (2.14) are good approximations only when the overall event rate is below 10% or $|\gamma_X| < 0.75$, i.e., $0.5 < \Gamma_X < 2$.

2.3 An Alternative Parameterization for the Association between U and X

In Sections 2.1 and 2.2, we parameterized the association between the unmeasured confounder U and the exposure indicator X in terms of the conditional distribution of U given X. An alternative parameterization, as used by Rosenbaum and Rubin (1983a) and Rosenbaum (1995, Chapter 4), specifies the effect of U on X. There is a simple relationship between these two types of parameterizations when U is binary. Specifically, let ϕ and P be the overall prevalences of the exposure and of the unmeasured confounder, respectively, and let Ψ be the probability ratio of being exposed associated with the unmeasured confounder, i.e., $\phi = \Pr(X = 1)$, $P = \Pr(U = 1)$, and $\Psi = \Pr(X = 1 \mid U = 1)/\Pr(X = 1 \mid U = 0)$. Note that Ψ measures directly the degree of departure from randomization due to the unmeasured confounder. By Bayes' theorem,

$$\Psi = \frac{P_1/P}{(1 - P_1)/(1 - P)},$$

which implies that

$$P_1 = \frac{\Psi P}{1 - P + \Psi P}. \tag{2.20}$$

On the other hand, because $P = P_1 \phi + P_0(1 - \phi)$, we have

$$P_0 = \frac{P - \phi P_1}{1 - \phi}. \tag{2.21}$$

By replacing P_1 and P_0 in (2.9) with (2.20) and (2.21) and by replacing P_1 in (2.21) with expression (2.20), we obtain an expression that relates R^* to R through Γ_0, Γ_1, Ψ, P, and ϕ.

We may also measure the effect of U on X by the odds ratio, say, Q. Clearly,

$$Q = \frac{P_1(1 - P_0)}{P_0(1 - P_1)}. \tag{2.22}$$

Replacing P_0 in (2.22) by expression (2.21), we can express P_1 as a function of P, ϕ, and Q by solving a quadratic equation. The resulting expression for P_1, together with (2.21), enables us to relate R^* and R through Γ_0, Γ_1, Q, P, and ϕ. This formulation is essentially the same as that of Rosenbaum and Rubin (1983a). In view of this connection, the proposed methodology may be regarded as a generalization of the pioneering work by Rosenbaum and Rubin. It seems preferable to use the probability ratio formulation rather than the odds ratio formulation for the effect of U on X not only because the former leads to simpler formulas but also because the probability ratio is more relevant than the odds ratio, especially when the exposure is not rare enough to ensure adequate approximation of the probability ratio by the odds ratio.

The parameter ϕ may be estimated from the current study or from an external data source, the resulting estimator being denoted by $\hat{\phi}$. With the restriction that $\Gamma_0 = \Gamma_1 = \Psi$ or $\Gamma_0 = \Gamma_1 = Q$, the results of the sensitivity analysis can be summarized in a concise manner, especially if the results are insensitive to the variation of P over a broad range of plausible values. The estimation of the variance of $\hat{\beta}$ or $\hat{R}$ is more complicated under the specification of Ψ or Q than under the direct specification of (P_0, P_1) if one wishes to adjust for the extra random variation due to the estimation of ϕ. Nonetheless, it is straightforward to establish the asymptotic joint distribution of $(\hat{\beta}^*, \hat{\phi})$ by the multivariate central limit theorem and the law of large numbers; the asymptotic normality and variance of $\hat{\beta}$ then follow from the delta method. The variation of $\hat{\phi}$ is negligible for large studies.

3. Survival Time

Let T denote the survival time or failure time of interest, and let X, $\mathbf{Z}$, and U be as defined at the start of Section 2. The proportional hazards regression (Cox, 1972) specifies that the hazard functions of T conditional on the sets of covariates $(X, \mathbf{Z}, U)$ and $(X, \mathbf{Z})$ are, respectively,

$$\lambda(t \mid X, \mathbf{Z}, U) = \lambda_0(t) \exp\left(\beta X + \gamma_X U + \boldsymbol{\theta}'\mathbf{Z}\right) \tag{3.1}$$

and

$$\lambda(t \mid X, \mathbf{Z}) = \lambda_0^*(t) \exp\left(\beta^* X + \boldsymbol{\theta}^{*'}\mathbf{Z}\right), \tag{3.2}$$

where $\lambda_0(\cdot)$ and $\lambda_0^*(\cdot)$ are arbitrary baseline hazard functions, and $(\beta, \gamma_0, \gamma_1, \boldsymbol{\theta})$ and $(\beta^*, \boldsymbol{\theta}^*)$ are unknown regression parameters. Models (3.1) and (3.2) parallel (2.1) and (2.2) as well as (2.3) and (2.16). As in Section 2, we want to ascertain the relationship between β^* and β.

Again, let $F(u \mid X, \mathbf{Z})$ be the distribution function of U given X and $\mathbf{Z}$. Also, let $f(t \mid \cdot)$ and $S(t \mid \cdot)$ denote, respectively, the conditional density and survival functions of T. By elementary probability arguments,

$$\lambda(t \mid X, \mathbf{Z}) = \frac{f(t \mid X, \mathbf{Z})}{S(t \mid X, \mathbf{Z})} = \frac{\int_{-\infty}^{\infty} f(t \mid X, \mathbf{Z}, u) dF(u \mid X, \mathbf{Z})}{\int_{-\infty}^{\infty} S(t \mid X, \mathbf{Z}, u) dF(u \mid X, \mathbf{Z})}. \tag{3.3}$$

Under model (3.1),

$$\int_{-\infty}^{\infty} f(t \mid X, \mathbf{Z}, u) dF(u \mid X, \mathbf{Z}) = \int_{-\infty}^{\infty} \lambda_0(t) e^{\beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}}$$
$$\times \exp\left\{-\Lambda_0(t) e^{\beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}}\right\} dF(u \mid X, \mathbf{Z}),$$

$$\int_{-\infty}^{\infty} S(t \mid X, \mathbf{Z}, u) dF(u \mid X, \mathbf{Z}) = \int_{-\infty}^{\infty} \exp\left\{-\Lambda_0(t) e^{\beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}}\right\} dF(u \mid X, \mathbf{Z}),$$

where $\Lambda_0(t) = \int_0^t \lambda_0(s) ds$. Thus, (3.3) becomes

$$\lambda(t \mid X, \mathbf{Z}) = \lambda_0(t) \exp\left(\beta X + \boldsymbol{\theta}'\mathbf{Z}\right) h(t; X, \mathbf{Z}), \tag{3.4}$$

where

$$h(t; X, \mathbf{Z}) = \frac{\int_{-\infty}^{\infty} e^{\gamma_X u} \exp\{-\Lambda_0(t) e^{\beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}}\} dF(u \mid X, \mathbf{Z})}{\int_{-\infty}^{\infty} \exp\left\{-\Lambda_0(t) e^{\beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}}\right\} dF(u \mid X, \mathbf{Z})}.$$

We evaluate the right side of (3.4) for the binary and normal U.

Binary confounder. If $F(u \mid X, \mathbf{Z})$ is a Bernoulli distribution with success probability $P_{X,\mathbf{Z}}$, then

$$h(t; X, \mathbf{Z}) = \frac{e^{\gamma_X} \exp\left\{-\Lambda_0(t)e^{\beta X + \gamma_X + \boldsymbol{\theta}'\mathbf{Z}}\right\} P_{X,\mathbf{Z}} + \exp\left\{-\Lambda_0(t)e^{\beta X + \boldsymbol{\theta}'\mathbf{Z}}\right\}(1 - P_{X,\mathbf{Z}})}{\exp\left\{-\Lambda_0(t)e^{\beta X + \gamma_X + \boldsymbol{\theta}'\mathbf{Z}}\right\} P_{X,\mathbf{Z}} + \exp\left\{-\Lambda_0(t)e^{\beta X + \boldsymbol{\theta}'\mathbf{Z}}\right\}(1 - P_{X,\mathbf{Z}})}. \quad (3.5)$$

To make the right side of (3.5) more tractable, we assume that the event is rare or the effect of U on T is small, which implies that $\Lambda_0(t)$ or $|\gamma_X|$ is small. Then (3.5) simplifies to $e^{\gamma_X} P_{X,\mathbf{Z}} + (1 - P_{X,\mathbf{Z}})$, which entails that (3.4) is

$$\lambda(t \mid X, \mathbf{Z}) \approx \lambda_0(t) \exp\left(\beta X + \boldsymbol{\theta}'\mathbf{Z}\right) \left\{e^{\gamma_X} P_{X,\mathbf{Z}} + (1 - P_{X,\mathbf{Z}})\right\}. \quad (3.6)$$

If U is conditionally independent of $\mathbf{Z}$ given X, then (3.6) is equivalent to

$$\lambda(t \mid X, \mathbf{Z}) \approx \lambda_0(t) \left\{e^{\gamma_0} P_0 + (1 - P_0)\right\} \exp\left[\left\{\beta + \log \frac{e^{\gamma_1} P_1 + (1 - P_1)}{e^{\gamma_0} P_0 + (1 - P_0)}\right\} X + \boldsymbol{\theta}'\mathbf{Z}\right]. \quad (3.7)$$

In view of (3.2) and (3.7),

$$\beta \approx \beta^* - \log \frac{e^{\gamma_1} P_1 + (1 - P_1)}{e^{\gamma_0} P_0 + (1 - P_0)}, \quad (3.8)$$

which takes the same form as formula (2.8).

Normal confounder. For the normally distributed U, it can be shown that

$$h(t; X, \mathbf{Z}) = \exp\left(\gamma_X \mu_{X,\mathbf{Z}} + 0.5\gamma_X^2\right) \frac{\int_{-\infty}^{\infty} \exp\left\{-\Lambda_0(t)e^{\gamma_X^2 + \beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}}\right\} \exp\left\{-\frac{(u - \mu_{X,\mathbf{z}})^2}{2}\right\} du}{\int_{-\infty}^{\infty} \exp\left\{-\Lambda_0(t)e^{\beta X + \gamma_X u + \boldsymbol{\theta}'\mathbf{Z}}\right\} \exp\left\{-\frac{(u - \mu_{X,\mathbf{z}})^2}{2}\right\} du}, \quad (3.9)$$

which can be approximated by $\exp(\gamma_X \mu_{X,\mathbf{Z}} + 0.5\gamma_X^2)$ if $\Lambda_0(t)$ or $|\gamma_X|$ is small. Then

$$\lambda(t \mid X, \mathbf{Z}) \approx \lambda_0(t) \exp\left(\beta X + \boldsymbol{\theta}'\mathbf{Z} + \gamma_X \mu_{X,\mathbf{Z}} + 0.5\gamma_X^2\right), \quad (3.10)$$

which is similar to (2.10). Under $\mu_{X,\mathbf{Z}} = \mu_X$, equation (3.10) entails that

$$\beta \approx \beta^* - \left\{(\gamma_1 \mu_1 - \gamma_0 \mu_0) + 0.5\left(\gamma_1^2 - \gamma_0^2\right)\right\}. \quad (3.11)$$

If $\gamma_0 = \gamma_1 = \gamma$ and $\mu_{X,\mathbf{Z}} = \mu_X + q(\mathbf{Z})$, then

$$\beta \approx \beta^* - \gamma\delta, \quad (3.12)$$

where again $\delta = \mu_1 - \mu_0$. Formulas (3.11) and (3.12) parallel (2.11) and (2.13).

Under model (3.1) with rare events, most of the data information about the event is contained in the event indicator rather than the event time. Because the event indicator is binary with approximate mean (2.1), it is not surprising that equations (3.8), (3.11), and (3.12) are the same as (2.8), (2.11), and (2.12).

The sensitivity analysis for survival time data can be performed in the same manner as that of binary outcome discussed in Section 2. Given the sensitivity parameters $(\gamma_X, P_{X,\mathbf{Z}})$ and $(\gamma_X, \mu_{X,\mathbf{Z}})$, the partial likelihood theory can be applied to models (3.6) and (3.10), respectively, to make inferences about β, as is discussed further in Section 5. If U is conditionally independent of $\mathbf{Z}$ given X, one may estimate β^* by fitting model (3.2) with any existing Cox regression software and then use formula (3.8) and (3.11) to conduct the sensitivity analysis. If $\gamma_0 = \gamma_1$ and $\mu_{X,\mathbf{Z}} = \mu_X + q(\mathbf{Z})$, then formula (3.12) may be used.

As in the setting of logistic regression, it is important to investigate how accurate the approximations (3.8), (3.11), and (3.12) are. In general, model (3.2) does not hold exactly because, as is evident from (3.5) and (3.9), the factor $h(t; X, \mathbf{Z})$ in (3.4) is a complicated function of t that depends on X and $\mathbf{Z}$, which implies that the right side of (3.4) does not reduce to that of (3.2). In particular, the hazard functions for the exposed and unexposed are not proportional over time. If one fits model (3.2) to the observable data, then the maximum partial likelihood estimator of β^*, denoted by $\hat{\beta}^*$, will converge to some well-defined limit, say, $\tilde{\beta}^*$, which represents an average over time of the log hazard ratios between the exposed and the unexposed (Struthers and Kalbfleisch, 1986; Lin and Wei, 1989). The question then arises as to how different $\tilde{\beta}^*$ is from β^* used in formulas (3.8), (3.11), and (3.12).

To answer the above question, we conducted a series of Monte Carlo studies similar to that of Section 2.2. The survival times were generated from model (3.1) with $\lambda_0(\cdot) = 1$, $\beta = \theta = 1$, $\gamma_0 = \gamma_1$, X taking the value 1 for 20% of the subjects and Z being normal with mean X and variance 1. The censoring times were generated from the uniform $(0, \tau)$ distribution, where τ was chosen such that a desired percentage of the survival times were censored. Again, there were 1000 simulation samples, each with 1000 subjects. For each sample, we calculated the maximum partial likelihood estimate $\hat{\beta}^*$ under model (3.2) and replaced β^* in formula (3.8) or (3.12) by $\hat{\beta}^*$ to yield $\hat{\beta}$. In the case of binary U, we set $P_0 = 0.2$ and $P_1 = 0.9$. Under $\Gamma = (2, 3, 4)$, the Monte Carlo estimates for the sampling means of $\hat{\beta}$ were approximately $(1.01, 1.01, 1.02)$ with 90% censoring and $(1.00, 0.98, 0.97)$ with 50% censoring. Similar results were obtained for other values of P_0 and P_1. In the case of normal U, we set $\mu_0 = 0$ and $\mu_1 = 1$ and created 90% censoring. The sampling means of $\hat{\beta}$ were found to be approximately $(0.99, 0.95, 0.88, 0.79)$ under $\gamma = (0.25, 0.5, 0.75, 1)$. The bias of $\hat{\beta}$ increased slightly when μ_1 was changed from 1 to 2 and decreased as censoring became heavier.

These numerical results suggest that, for binary U, the maximum partial likelihood estimator $\hat{\beta}^*$ under model (3.2) may be used in formula (3.8) to yield an approximately unbiased estimator of β given the sensitivity parameters, whereas for normal U, formulas (3.11) and (3.12) are good approximations only when the censoring percentage is greater than 90% or $|\gamma_X| < 0.75$, i.e., $0.5 < \Gamma_X < 2$. The variance of $\hat{\beta}^*$ may be estimated by the partial likelihood method or, more appropriately, by the sandwich estimator of Lin and Wei (1989), which is available in most software packages. For correlated survival times, the robust variance estimator can be found in Lin (1994) and is also available in software packages.

4. Real Examples

We now apply the methods developed in Sections 2 and 3 to two recently published medical studies. The first one is a case–control study investigating the potential role of appetite-suppressant drugs in the development of primary pulmonary hypertension (Abenhaim et al., 1996), and the second one is a prospective cohort study examining the association between the use of right-heart catheterization during the first 24 hours of care in the intensive care unit and subsequent survival (Connors et al., 1996). Both studies aroused considerable controversy because they were observational studies that raised concerns about the safety of commonly used interventions.

4.1 *Primary Pulmonary Hypertension Study*

Primary pulmonary hypertension (PPH) is a rare, often fatal disease with unknown causes. In the early 1990s, PPH developed in a cluster of French patients exposed to derivatives of fenfluramine in appetite suppressants used for weight control. This phenomenon prompted a major European case–control study that evaluated 95 patients with PPH and 355 controls matched individually on gender and age. The investigators used conditional logistic regression to study the association between the exposure to appetite suppressants and the development of PPH after adjustment for weight-related confounding variables and other suspected risk factors. With the adjustment, the odds ratio of PPH among those who had taken appetite suppressants for a total of more than 3 months was estimated at 23.1, with a 95% confidence interval of $(6.9, 77.7)$.

Because the factors leading to the development of PPH remain enigmatic, it is worthwhile to assess the sensitivity of the findings of this case–control study to the assumptions about unmeasured confounders. To this end, we display in Table 1 the point and interval estimates for the odds ratio of interest after adjustment for an unmeasured binary confounder according to formula (2.9) with $\Gamma_0 = \Gamma_1 = \Gamma$. Here, P_1 and P_0 are, respectively, the prevalences of the unmeasured confounder among those who had used appetite suppressants for more than 3 months and among those who had never used them, and Γ is the odds ratio of PPH associated with the unmeasured confounder. We display only the estimates for $P_1 \geq P_0$. The estimates will be inflated upward if $P_1 < P_0$. In Table 1, the point estimates are always much higher than 1, whereas the lower limits of the 95% confidence intervals are greater than 1, except for two entries under $\Gamma = 10$. Recall that the adjustment factor A is bounded above by Γ. Thus, regardless of the values of P_0 and P_1, the point estimate will remain above 1 as long as Γ is less than 23.1, and the lower confidence limit will be greater than 1 if Γ is below 6.9. When Γ is larger than 6.9, the lower confidence limit may drop below 1 if P_0 is close to 0 and P_1 is close to 1.

It is also informative to assess how the adjustment for an unmeasured normal confounder might affect the estimates. Formula (2.14) implies that, under $\delta = 1$, the point or interval estimate adjusting for the normal confounder will remain above 1 as long as Γ is less than the unadjusted estimate. Under $\delta = 2$, the point estimate and the lower limit of the 95% confidence interval will remain above 1 if Γ is less than 4.81 and 2.63, respectively.

Table 1
The point estimates and 95% confidence intervals for the odds ratio of
PPH associated with long-term use of appetite suppressant drugs with
adjustment for an unmeasured binary confounder of specified properties

Γ	P_1	P_0 0.0	0.1	0.2	0.3	0.4	0.5
2	0.0	23.1 $(6.9, 77.7)$					
	0.2	19.3 $(5.8, 64.8)$	21.2 $(6.3, 71.2)$	23.1 $(6.9, 77.7)$			
	0.4	16.5 $(4.9, 55.5)$	18.2 $(5.4, 61.1)$	19.8 $(5.9, 66.6)$	21.5 $(6.4, 72.2)$	23.1 $(6.9, 77.7)$	
	0.6	14.4 $(4.3, 48.6)$	15.9 $(4.7, 53.4)$	17.3 $(5.2, 58.3)$	18.8 $(5.6, 63.1)$	20.2 $(6.0, 68.0)$	21.7 $(6.5, 72.8)$
	0.8	12.8 $(3.8, 43.2)$	14.1 $(4.2, 47.5)$	15.4 $(4.6, 51.8)$	16.7 $(5.0, 56.1)$	18.0 $(5.4, 60.4)$	19.3 $(5.8, 64.8)$
	1.0	11.6 $(3.5, 38.9)$	12.7 $(3.8, 42.7)$	13.9 $(4.1, 46.6)$	15.0 $(4.5, 50.5)$	16.2 $(4.8, 54.4)$	17.3 $(5.2, 58.3)$
6	0.0	23.1 $(6.9, 77.7)$					
	0.2	11.6 $(3.5, 38.9)$	17.3 $(5.2, 58.3)$	23.1 $(6.9, 77.7)$			
	0.4	7.7 $(2.3, 25.9)$	11.6 $(3.5, 38.9)$	15.4 $(4.6, 51.8)$	19.3 $(5.8, 64.8)$	23.1 $(6.9, 77.7)$	
	0.6	5.8 $(1.7, 19.4)$	8.7 $(2.6, 29.1)$	11.6 $(3.5, 38.9)$	14.4 $(4.3, 48.6)$	17.3 $(5.2, 58.3)$	20.2 $(6.0, 68.0)$
	0.8	4.6 $(1.4, 15.5)$	6.9 $(2.1, 23.3)$	9.2 $(2.8, 31.1)$	11.6 $(3.5, 38.9)$	13.9 $(4.1, 46.6)$	16.2 $(4.8, 54.4)$
	1.0	3.9 $(1.2, 13.0)$	5.8 $(1.7, 19.4)$	7.7 $(2.3, 25.9)$	9.6 $(2.9, 32.4)$	11.6 $(3.5, 38.9)$	13.5 $(4.0, 45.3)$
10	0.0	23.1 $(6.9, 77.7)$					
	0.2	8.3 $(2.5, 27.8)$	15.7 $(4.7, 52.7)$	23.1 $(6.9, 77.7)$			
	0.4	5.0 $(1.5, 16.9)$	9.5 $(2.9, 32.1)$	14.1 $(4.2, 47.3)$	18.6 $(5.6, 62.5)$	23.1 $(6.9, 77.7)$	
	0.6	3.6 $(1.1, 12.1)$	6.9 $(2.0, 23.1)$	10.1 $(3.0, 34.0)$	13.4 $(4.0, 44.9)$	16.6 $(5.0, 55.8)$	19.9 $(5.9, 66.8)$
	0.8	2.8 $(0.8, 9.5)$	5.4 $(1.6, 18.0)$	7.9 $(2.4, 26.5)$	10.4 $(3.1, 35.1)$	13.0 $(3.9, 43.6)$	15.5 $(4.6, 52.1)$
	1.0	2.3 $(0.7, 7.8)$	4.4 $(1.3, 14.8)$	6.5 $(1.9, 21.8)$	8.5 $(2.6, 28.7)$	10.6 $(3.2, 35.7)$	12.7 $(3.8, 42.7)$

The results of this sensitivity analysis strongly support the investigators' conclusion that the long-term use of appetite-suppressant drugs increased the risk of PPH. It should be noted that the type of sensitivity analyses used here does not and cannot account for potential detection and recall biases, which the investigators judged to be minimal in this study.

4.2 Right-Heart Catheterization Study

Many cardiologists and intensive-care physicians believe that the direct measurement of cardiac function provided by right-heart catheterization (RHC) is necessary to guide therapy for certain critically ill patients and that such management leads to better outcomes. This popular belief is so strong that it has prevented the conduct of a randomized clinical trial even though the benefit of RHC has never been demonstrated in a clinical trial. Connors et al. (1996) reported a prospective cohort study designed to assess the effectiveness of RHC in the initial care of critically ill patients. The study involved 5735 patients receiving care in the intensive care unit for one of nine prespecified disease categories. The investigators did an exceptionally thorough analysis of their data. Not only did they attempt to identify all the major risk factors, but they also constructed a propensity score (Rosenbaum and Rubin, 1983b) for RHC to adjust for treatment selection bias. Using a

proportional hazards model that included the propensity score as well as nine risk factors, they estimated the hazard ratio or relative hazard of death for patients managed with RHC at 1.21 with a 95% confidence interval of $(1.09, 1.25)$. It is interesting to assess the sensitivity of this unexpected finding to potential unknown confounders.

Table 2 shows the point and interval estimates for the relative hazard of death associated with RHC after adjustment for an unmeasured binary confounder according to formula (3.8) with $\gamma_0 = \gamma_1 \approx 1.1$, i.e., $\Gamma_0 = \Gamma_1 = \Gamma = 3$. Here, P_1 and P_0 are, respectively, the prevalences of the unmeasured confounder among those who were managed with RHC and among those who were not, and Γ is the relative hazard of death for the unmeasured confounder. As shown in the table, the point estimates for the relative hazard of death are no larger than 1.01 if $P_1 - P_0 \geq 0.2$, and the lower limits of the 95% confidence intervals drop below 1 if $P_1 - P_0 \geq 0.1$. Furthermore, the point estimates are no larger than 0.81 if $(P_0, P_1) = (0, 0.3)$, $(0.1, 0.4)$, $(0.2, 0.6)$, $(0.3, 0.7)$, $(0.4, 0.9)$, or $(0.5, 1.0)$, and the upper confidence limits are about 0.8 or smaller if $P_1 - P_0 > 0.5$.

Table 3 displays the estimates for the relative hazard of death associated with RHC under the alternative parameterization described in Section 2.3. Here, Ψ is the probability ratio of RHC use associated with the unmeasured confounder, and Γ is again the relative hazard of death for the unmeasured confounder. Under $\Psi = \Gamma = 2$, the point estimates for the relative hazard of death are very close to the null value of 1 if $0.2 \leq P \leq 0.5$; under $\Psi = \Gamma = 3$, the estimates lie between 0.73 and 0.82 if $0.1 \leq P \leq 0.5$. In other words, over a broad range of plausible values for the prevalence of the unmeasured confounder, ranging from 0.1 to 0.5, a true relative hazard of 0.8 could be misrepresented as 1.21 if an unmeasured confounder increased both the hazard of death and the probability of RHC three-fold, and a true relative hazard of 1.0 could be misrepresented as 1.21 if an unmeasured confounder increased both the hazard of death and the probability of RHC two-fold.

Connors et al. (1996) also performed a sensitivity analysis. They stated that "the covariate missing from the propensity score would have to increase the risk of death 6-fold and increase the probability of RHC 6-fold for a true relative hazard of 0.80 to be misrepresented as a relative hazard of 1.21. In addition, a true relative hazard of 1.0 could be misrepresented as 1.21 if a missing

Table 2

The point estimates and 95% confidence intervals for the relative hazard of death associated with RHC use with adjustment for an unmeasured binary confounder having relative hazard of death of 3

			P_0			
P_1	0.0	0.1	0.2	0.3	0.4	0.5
0.0	1.21					
	$(1.09, 1.25)$					
0.1	1.01	1.21				
	$(0.91, 1.04)$	$(1.09, 1.25)$				
0.2	0.86	1.04	1.21			
	$(0.78, 0.89)$	$(0.93, 1.07)$	$(1.09, 1.25)$			
0.3	0.76	0.91	1.06	1.21		
	$(0.68, 0.78)$	$(0.82, 0.94)$	$(0.95, 1.09)$	$(1.09, 1.25)$		
0.4	0.67	0.81	0.94	1.08	1.21	
	$(0.61, 0.69)$	$(0.73, 0.83)$	$(0.85, 0.97)$	$(0.97, 1.11)$	$(1.09, 1.25)$	
0.5	0.60	0.73	0.85	0.97	1.09	1.21
	$(0.55, 0.62)$	$(0.65, 0.75)$	$(0.76, 0.88)$	$(0.87, 1.00)$	$(0.98, 1.12)$	$(1.09, 1.25)$
0.6	0.55	0.66	0.77	0.88	0.99	1.10
	$(0.50, 0.57)$	$(0.59, 0.68)$	$(0.69, 0.80)$	$(0.79, 0.91)$	$(0.89, 1.02)$	$(0.99, 1.14)$
0.7	0.50	0.60	0.71	0.81	0.91	1.01
	$(0.45, 0.52)$	$(0.55, 0.62)$	$(0.64, 0.73)$	$(0.73, 0.83)$	$(0.82, 0.94)$	$(0.91, 1.04)$
0.8	0.47	0.56	0.65	0.74	0.84	0.93
	$(0.42, 0.48)$	$(0.50, 0.58)$	$(0.59, 0.67)$	$(0.67, 0.77)$	$(0.75, 0.87)$	$(0.84, 0.96)$
0.9	0.43	0.52	0.60	0.69	0.78	0.86
	$(0.39, 0.45)$	$(0.47, 0.54)$	$(0.55, 0.62)$	$(0.62, 0.71)$	$(0.70, 0.80)$	$(0.78, 0.89)$
1.0	0.40	0.48	0.56	0.65	0.73	0.81
	$(0.36, 0.42)$	$(0.44, 0.50)$	$(0.51, 0.58)$	$(0.58, 0.67)$	$(0.65, 0.75)$	$(0.73, 0.83)$

Note: P_1 and P_0 are the prevalences of the unmeasured confounder among the RHC users and nonusers, respectively.

<div align="center">

Table 3

The prevalences of the unmeasured confounder among the users and nonusers of
RHC, P_1, and P_0 and the corresponding point estimates and 95% confidence intervals
for the relative hazard of death associated with RHC, $\hat{R}$, and 95% CI under a direct
parameterization for the effect of an unmeasured binary confounder on the use of RHC

</div>

		$\Psi = \Gamma = 2$				$\Psi = \Gamma = 3$		
P	P_1	P_0	$\hat{R}$	95% CI	P_1	P_0	$\hat{R}$	95% CI
0.1	0.18	0.05	1.07	$(0.97, 1.11)$	0.25	0.01	0.82	$(0.74, 0.85)$
0.2	0.33	0.12	1.01	$(0.91, 1.05)$	0.43	0.06	0.73	$(0.66, 0.75)$
0.3	0.46	0.20	0.99	$(0.90, 1.03)$	0.56	0.14	0.73	$(0.66, 0.75)$
0.4	0.57	0.29	1.00	$(0.90, 1.03)$	0.67	0.24	0.76	$(0.69, 0.79)$
0.5	0.67	0.40	1.01	$(0.91, 1.05)$	0.75	0.35	0.82	$(0.74, 0.85)$
0.6	0.75	0.51	1.04	$(0.94, 1.08)$	0.82	0.47	0.89	$(0.80, 0.92)$
0.7	0.82	0.62	1.08	$(0.97, 1.11)$	0.88	0.59	0.96	$(0.87, 0.99)$
0.8	0.89	0.75	1.12	$(1.01, 1.16)$	0.92	0.72	1.04	$(0.94, 1.08)$
0.9	0.95	0.87	1.16	$(1.05, 1.20)$	0.96	0.86	1.12	$(1.01, 1.16)$

Note: P is the overall prevalence of the unmeasured confounder; Ψ and Γ are, respectively, the probability ratio of RHC and relative hazard of death associated with the unmeasured confounder. The observed prevalence of RHC, $\hat{\phi}$, was 0.38, with an estimated standard error of 0.006. We did not adjust for the variation of $\hat{\phi}$ in constructing the 95% confidence intervals, but the variation is too small to make an appreciable difference.

covariate that simultaneously increased the risk of death 3-fold and the probability of RHC 3-fold was not accounted for in the propensity score." These results appear to differ from those described in the preceding paragraph. There are several reasons for these discrepancies. First, because of the lack of sensitivity analysis methods for Cox regression, Connors et al. obtained their results by applying the method of Rosenbaum and Rubin (1983a) to the binary outcome of 30-day mortality rather than patient survival time. The point estimate for the odds ratio of death as actually used in their sensitivity analysis was 1.29, which is slightly higher than the reported relative hazard estimate of 1.21. More important, the method of Rosenbaum and Rubin formulates the effect of the unmeasured confounder on treatment selection, which in this case is the use of RHC, in terms of the odds ratio, but Connors et al. inadvertently described this odds ratio as the probability ratio in their statement. Because RHC is a common procedure, an increase in the probability of RHC translates into a larger increase in the odds of RHC. Thus, misinterpreting the odds ratio of RHC as the probability ratio of RHC resulted in an overestimation of the effects of an unmeasured confounder that would be required to misrepresent a neutral or beneficial effect of RHC as harmful. Dr. Connors confirmed that "risk of death," "probability of RHC," and "relative hazard of 1.21" in their statement should read "odds of death," "odds of RHC," and "odds ratio of 1.29." Given this clarification, we performed an additional sensitivity analysis on the odds ratio estimate of 1.29 using the sensitivity parameters that were actually used by Conners et al. For $P = 0.5$, which is the overall prevalence implicitly assumed by Connors et al., we estimated the true odds ratios of death associated with RHC at 0.99 and 0.73 under $Q = \Gamma = 3$ and $Q = \Gamma = 6$, respectively, which are close to the values of 1 and 0.8 obtained by Connors et al. Here, Q is the odds ratio of RHC associated with the unmeasured confounder.

It should be pointed out that, because of the extraordinary efforts made by the investigators to identify and adjust for all the known risk factors, it is unlikely that there exists an unmeasured confounder with $Q = \Gamma = 3$ or 6 even though these were the values considered by the investigators in their sensitivity analysis. On the other hand, the lower limit of the observed 95% confidence interval was 1.09, which is not far above the null value of 1.0. Thus, the adjustment for an unmeasured confounder that has only modest effects on the use of RHC and death would drop the lower confidence limit below 1 or even 0.8.

The results of our sensitivity analysis provide additional insights into this important study and imply perhaps greater uncertainty about the role of RHC than those stated in the original report. The observed relative hazard of 1.21 with a 95% confidence interval of (1.09, 1.25) suggests a modest increase in the hazard of death due to RHC and could represent a true relative hazard of 1 or 0.8 if there existed an unmeasured confounder of moderate independent effects. Thus, this study did

not provide strong evidence for either a harmful or beneficial effect of RHC. It would be desirable to carry out a randomized controlled clinical trial, as recommended by Connors et al. (1996).

Incidentally, the censoring percentage in this study was approximately 65%. According to the simulation results reported at the end of Section 3, formula (3.8) is a good approximation under such a censoring level, whereas formulas (3.11) and (3.12) may not be. Thus, we did not perform a sensitivity analysis with an unmeasured normal confounder for this study.

5. Discussion and Further Results

Observational studies are subject to potential biases induced by unmeasured or unknown confounders. Many epidemiologists use a rule of thumb that the actual relative risk may be off by a factor of 2 because of unrecognized confounders. That rule comports well with the results shown in Tables 1–3.

If $\Gamma_0 = \Gamma_1$, then formula (2.9) takes the same form as formula (1) of Schlesselman (1978), which becomes formula (1) of Cornfield et al. (1959) under $R = 1$. Thus, Section 2 establishes that the basic relationships between the true and apparent associations derived by Cornfield and Schlesselman for the simple 2×2 table with an unmeasured binary confounder continue to hold when the log-linear/logistic regression models are used to adjust for the measured confounders that are conditionally independent of the unmeasured binary confounder. Section 3 further demonstrates that similar relationships hold for the Cox regression with censored survival data. The availability of simple and explicit formulas, such as (2.8) and (2.9), (2.11) to (2.14), (3.8), and (3.11) to (3.12), will facilitate the use of the powerful tool of sensitivity analysis by practitioners. It is straightforward to show that these formulas hold when the corresponding regression models are further stratified on some categorical variables provided that β, $\boldsymbol{\theta}$, and the sensitivity parameters remain the same among the strata.

Compared with the published methods, the proposed technique has the advantage of accommodating both discrete and continuous covariates. In observational studies, it is desirable and customary to collect information on all previously identified confounders and to adjust for them in the estimation of the exposure effect. The method of sensitivity analysis presented in this paper enables one to assess the residual confounding effects due to unmeasured or unknown covariates after adjusting for the effects of the measured covariates. Without the adjustment for the known and measured confounders, the sensitivity analysis would need to encompass much broader ranges of values for the sensitivity parameters and therefore become less informative.

One complication in assessing the sensitivity of the estimate of exposure effect that adjusts for covariates through nonlinear regression is that the reduced model that omits the unmeasured confounder does not hold exactly. However, the analytical and numerical investigations in Sections 2 and 3 demonstrated that the reduced logistic model (2.16) and proportional hazards model (3.2) provide reasonable approximations for rare events. The requirement of rare events is not a major limitation of the proposed methodology because the rarity of the event is often one of the main reasons for conducting observational studies, especially case–control studies. Furthermore, the numerical studies in Sections 2 and 3 indicated that, for unmeasured binary confounders, the approximations are adequate even if events are not rare.

Formulas (2.8) and (2.9), (2.11) and (2.12), (3.8), and (3.11) require the conditional independence of U and $\mathbf{Z}$ given X. This assumption may be satisfied if, e.g., U is a genetic trait and $\mathbf{Z}$ consists of environmental factors. To avoid this assumption, one can fit model (2.17) or (3.4) to the observable data by specifying the nature of dependence between U and $\mathbf{Z}$, that is, $P_{X,\mathbf{Z}}$ or $\mu_{X,\mathbf{Z}}$, together with γ_X. For rare events with binary or normal U, the likelihood score function and observed Fisher information matrix for $(\alpha, \beta, \boldsymbol{\theta})$ of model (2.17) are, respectively,

$$\dot{\mathbf{l}}(\alpha, \beta, \boldsymbol{\theta}) = \sum_{i=1}^{n} \{Y_i - \pi(X_i, \mathbf{Z}_i)\} \begin{bmatrix} 1 \\ X_i \\ \mathbf{Z}_i \end{bmatrix},$$

$$\ddot{\mathbf{l}}(\alpha, \beta, \boldsymbol{\theta}) = \sum_{i=1}^{n} \pi(X_i, \mathbf{Z}_i) \{1 - \pi(X_i, \mathbf{Z}_i)\} \begin{bmatrix} 1 & X_i & \mathbf{Z}_i' \\ X_i & X_i^2 & X_i \mathbf{Z}_i' \\ \mathbf{Z}_i & X_i \mathbf{Z}_i & \mathbf{Z}_i \mathbf{Z}_i' \end{bmatrix},$$

where

$$\pi(X, \mathbf{Z}) = \frac{\exp\{\alpha + \beta X + \boldsymbol{\theta}'\mathbf{Z} + g(X, \mathbf{Z})\}}{1 + \exp\{\alpha + \beta X + \boldsymbol{\theta}'\mathbf{Z} + g(X, \mathbf{Z})\}}.$$

The key difference between the current setting and standard logistic regression is that here $\pi(X, \mathbf{Z})$ involves the extra term $g(X, \mathbf{Z})$, but $g(X, \mathbf{Z})$ is free of $(\alpha, \beta, \boldsymbol{\theta})$. Let $(\hat{\alpha}, \hat{\beta}, \hat{\boldsymbol{\theta}})$ be the solution to $\dot{\mathbf{l}}(\alpha, \beta, \boldsymbol{\theta}) = 0$. For large n, $(\hat{\alpha}, \hat{\beta}, \hat{\boldsymbol{\theta}})$ is approximately normal with mean $(\alpha, \beta, \boldsymbol{\theta})$ and covariance matrix $\ddot{\mathbf{I}}^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\boldsymbol{\theta}})$. For rare events with binary or normal U, the factor $h(t; X, \mathbf{Z})$ in (3.4) does not depend on t or $(\beta, \boldsymbol{\theta})$; therefore, we can apply the partial-likelihood procedures for standard Cox regression to model (3.4) by replacing the exponential regression function $\exp(\beta X + \boldsymbol{\theta}'\mathbf{Z})$ involved in the score function and information matrix with $\exp(\beta X + \boldsymbol{\theta}'\mathbf{Z})h(t; X, \mathbf{Z})$. If events are not rare, then the maximum likelihood estimation of model (2.17) is complicated and the partial likelihood for model (3.4) intractable.

In applications, the direct fitting of models (2.17) and (3.4) is not very appealing for several reasons. First, it is difficult to specify the dependence between U and $\mathbf{Z}$ because U is unobserved; such a detailed specification leads to too many sensitivity parameters and consequently a less informative analysis. Second, the fitting of these two models requires the access to raw data. Third, because of the involvement of $g(X, \mathbf{Z})$ and $h(X, \mathbf{Z})$, standard software for logistic and Cox regression cannot be used directly.

It is of practical importance to ascertain what kind of approximations formulas (2.8) and (2.9), (2.11) and (2.12), (3.8), and (3.11) provide when U and $\mathbf{Z}$ are not conditionally independent given X. For rare events with binary U, (2.17) can be written as

$$\operatorname{logit}\{\Pr(Y = 1 \mid X, \mathbf{Z})\} \approx \left[\alpha + \log\left\{e^{\gamma_0}P_{0,\mathbf{Z}} + (1 - P_{0,\mathbf{Z}})\right\}\right]$$
$$+ \left\{\beta + \log \frac{e^{\gamma_1}P_{1,\mathbf{Z}} + (1 - P_{1,\mathbf{Z}})}{e^{\gamma_0}P_{0,\mathbf{Z}} + (1 - P_{0,\mathbf{Z}})}\right\}X + \boldsymbol{\theta}'\mathbf{Z}, \quad (5.1)$$

which reduces to (2.18) if $P_{X,\mathbf{Z}} = P_X$. Without the conditional independence, (5.1) does not reduce to (2.18), which implies that model (2.16) does not hold. If one still fits model (2.16) to the observable data $(Y_i, X_i, \mathbf{Z}_i)$ $(i = 1, \ldots, n)$, then, as mentioned in Section 2.2, the maximum likelihood estimator $(\hat{\alpha}^*, \hat{\beta}^*, \hat{\boldsymbol{\theta}}^*)$ converges to the constant vector, denoted by $(\tilde{\alpha}^*, \tilde{\beta}^*, \tilde{\boldsymbol{\theta}}^*)$, which minimizes the Kullback–Leibler distance between models (5.1) and (2.16). Comparison of the regression coefficients for X between these two models suggests that

$$\tilde{\beta}^* \leq \beta + \log \tilde{A},$$

or

$$\beta \geq \tilde{\beta}^* - \log \tilde{A}, \quad (5.2)$$

where

$$\tilde{A} = \max_{\mathbf{Z}} \frac{e^{\gamma_1}P_{1,\mathbf{Z}} + (1 - P_{1,\mathbf{Z}})}{e^{\gamma_0}P_{0,\mathbf{Z}} + (1 - P_{1,\mathbf{Z}})}.$$

Inequality (5.2) implies that the right side of (2.8) provides a lower bound for β if (P_1, P_0) in (2.8) is the combination of $(P_{1,\mathbf{Z}}, P_{0,\mathbf{Z}})$ that yields the maximum adjustment factor A among all possible values of $\mathbf{Z}$. Formulas (2.9), (2.11) to (2.12), (3.8), and (3.11) may be interpreted similarly.

In this paper, we focused on binary response and censored survival data because they are the two most common endpoints in observational studies, especially in biomedical applications. The approach taken here may also be applied to other outcome measures. It is worthwhile to discuss briefly the linear regression for continuous response variables. Let Y be a continuous response variable, and let X, $\mathbf{Z}$, and U be as defined in Sections 2 and 3. Suppose that Y is related to X, $\mathbf{Z}$, and U through the linear model

$$Y = \beta X + \gamma_X U + \boldsymbol{\theta}'\mathbf{Z} + \epsilon,$$

and that the model being fitted is

$$Y = \beta^* X + \boldsymbol{\theta}^{*\prime}\mathbf{Z} + \epsilon^*,$$

where $(\beta, \gamma_0, \gamma_1, \boldsymbol{\theta})$ and $(\beta^*, \boldsymbol{\theta}^*)$ are unknown regression parameters, and ϵ and ϵ^* are error terms with unspecified distributions. For survival data, it is common to set $Y = \log T$, which corresponds to the accelerated failure time model (Kalbfleisch and Prentice, 1980, pp. 32–34). It is easy to show that, under the conditional independence of U and $\mathbf{Z}$ given X,

$$\beta = \beta^* - (\gamma_1 m_1 - \gamma_0 m_0), \quad (5.3)$$

where m_1 and m_0 are the means of U for the exposed and the unexposed, respectively. If $\gamma_0 = \gamma_1 = \gamma$, then equation (5.3) reduces to

$$\beta = \beta^* - \gamma d, \tag{5.4}$$

where $d = m_1 - m_0$. Formula (5.4) does not require the conditional independence of U and $\mathbf{Z}$, but the effects of X and $\mathbf{Z}$ on the mean of U must be additive. Rosenbaum (1986) used a formula similar to (5.4) to study the effect of dropping out of high school on cognitive achievement test scores. For noncensored data, β^* can be estimated by the least-squared method. Draper and Smith (1981, Section 12) gave a general account of the bias in the least-squared estimator due to omission of covariates. The estimation of β^* for censored survival data has been studied by Tsiatis (1990); Wei, Ying, and Lin (1990); and Lai and Ying (1991).

The new method as well as all the published methods deal only with a single unmeasured confounder. The approach taken in Sections 2 and 3 can be generalized to the setting of multiple unmeasured confounders. One would then need to specify the joint distribution of the confounders as well as their joint effects on the outcome variable. A more practical alternative is to simply treat U as a composite score constructed from a set of unmeasured confounders. In this regard, a normal U may represent a linear combination of several confounders and a binary U the dichotomy of high risk versus low risk determined by multiple risk factors. The use of a single unmeasured confounder is a less severe restriction when all the known confounders are adjusted for through regression modeling.

This paper complements the current literature on misspecified nonlinear regression models. Gail, Wieand and Piantadosi (1984), Lagakos and Schoenfeld (1984), and Struthers and Kalbfleisch (1986) investigated the bias for the estimator of treatment effect in randomized experiments when relevant covariates are omitted from logistic and proportional hazards models. Our paper extends their results to nonrandomized studies in which the omitted covariates are not independent of treatment.

Acknowledgements

Résumé

L'article présente une approche générale qui permet dans une étude d'observation, d'évaluer la sensibilité, à de potentielles covariables non observées, de l'estimation et de l'intervalle de confiance correspondant à l'effet d'une exposition, après ajustement sur des variables connues. La méthode proposée suppose que l'effet réel de l'exposition puisse être modélisé par une régression qui incluera l'indicateur d'exposition ainsi que les covariables observées et non observées. On peut alors utiliser un modèle dit "réduit", car il ne prend pas en compte les covariables non observées, pour obtenir une inférence sur l'effet réel de l'exposition, en spécifiant la distribution de la covariable omise pour le groupe exposé et le groupe non exposé ainsi que la relation entre la covariable non observée et l'événement d'intérêt. Sous certaines conditions, il existe une relation algébrique simple entre l'effet réel de l'exposition sous le modèle complet et l'effet observé de l'exposition sous le modèle réduit. On peut alors estimer l'effet réel de l'exposition par un simple ajustement de l'estimation et de l'intervalle de confiance de l'effet observé préalablement obtenu par les méthodes de calculs classiques ou par les données de la littérature. La méthode proposée s'applique aussi bien aux données binaires qu'aux données censurées, elle est adaptée à tous les types d'étude et peut prendre en compte des variables oubliées discrètes ou distribuées selon une loi normale. Nous présentons l'application de ces méthodes à deux grandes études médicales.

References

Abenhaim, L., Moride, Y., Brenot, F., et al. (1996). Appetite-suppressant drugs and the risk of primary pulmonary hypertension. *The New England Journal of Medicine* **335,** 609–616.

Bross, I. D. J. (1966). Spurious effects from an extraneous variable. *Journal of Chronic Diseases* **19,** 637–647.

Connors, A. F., Speroff, T., Dawson, N. V., et al. (1996). The effectiveness of right heart catheterization in the initial care of critically ill patients. *Journal of the American Medical Association* **276,** 889–897.

Cornfield, J., Haenszel, W., Hammond, E. C., Lilienfeld, A. M., Shimkin, M. B., and Wynder, E. L. (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *Journal of the National Cancer Institute* **22,** 173–203.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B* **34,** 187–220.

Draper, N. and Smith, H. (1981). *Applied Regression Analysis.* 2nd Edition. New York: Wiley.

Fisher, R. A. (1958). Lung cancer and cigarettes? *Nature* **182,** 108.

Gail, M. H., Wieand, S., and Piantadosi, S. (1984). Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika* **71,** 431–444.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: Wiley.

Lagakos, S. W. and Schoenfeld, D. A. (1984). Properties of proportional-hazards score tests under misspecified regression models. *Biometrics* **40,** 1037–1048.

Lai, T. L. and Ying, Z. (1991). Large sample theory of a modified Buckley–James estimator for regression analysis with censored data. *The Annals of Statistics* **19,** 1370–1402.

Lin, D. Y. (1994). Cox regression analysis of multivariate failure time data: The marginal approach. *Statistics in Medicine* **13,** 2233–2247.

Lin, D. Y. and Wei, L. J. (1989). The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association* **84,** 1074–1078.

Rosenbaum, P. R. (1986). Dropping out of high school in the United States: An observational study. *Journal of Educational Statistics* **11,** 207–224.

Rosenbaum, P. R. (1995). *Observational Studies.* New York: Springer-Verlag.

Rosenbaum, P. R. and Rubin, D. B. (1983a). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society, Series B* **11,** 212–218.

Rosenbaum, P. R. and Rubin, D. B. (1983b). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70,** 41–55.

Schlesselman, J. J. (1978). Assessing effects of confounding variables. *American Journal of Epidemiology* **108,** 3–8.

Struthers, C. A. and Kalbfleisch, J. D. (1986). Misspecified proportional hazard models. *Biometrika* **73,** 363–369.

Tsiatis, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *The Annals of Statistics* **18,** 354–372.

Wei, L. J., Ying, Z., and Lin, D. Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika* **77,** 845–851.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* **50,** 1–25.