University of Pennsylvania
UPenn Biostatistics Working Papers

Year 2006 *Paper* 5

Sensitivity of the Hazard Ratio to Non-Ignorable Treatment Assignment in an Observational Study

Nandita Mitra* Daniel F. Heitjan†

*University of Pennsylvania, nmitra@cceb.upenn.edu

†University of Pennsylvania, dheitjan@cceb.upenn.edu

Sensitivity of the Hazard Ratio to Non-Ignorable Treatment Assignment in an Observational Study

Nandita Mitra and Daniel F. Heitjan

Abstract

In non-randomized studies, estimation of treatment effects generally requires adjustment for imbalances in observed covariates. One such method, based on the propensity score, is useful in many applications but may be biased when the assumption of strongly ignorable treatment assignment is violated. Because it is not possible to evaluate this assumption from the data, it is advisable to assess the sensitivity of conclusions to violations of strong ignorability. Lin et al [1] have implemented this idea by investigating how an unmeasured covariate may affect the conclusions of an observational study. We extend their method to assess sensitivity of the treatment hazard ratio to hidden bias under a range of covariate distributions. We derive simple formulas for approximating the true from the apparent treatment hazard ratio estimated under a specific survival model, and assess the validity of these formulas in simulation studies. We demonstrate the method in an analysis of SEER-Medicare data on the effects of chemotherapy in elderly colon cancer patients.

Sensitivity of the Hazard Ratio to Non-Ignorable Treatment Assignment in an Observational Study

Nandita Mitra, Ph.D.*
University of Pennsylvania School of Medicine
Department of Biostatistics and Epidemiology
612 Blockley Hall, 423 Guardian Drive
Philadelphia, PA 19104-6021
Email: nmitra@cceb.upenn.edu
Voice: 215-573-4467
Fax: 215-573-4865

Daniel F. Heitjan, Ph.D.
University of Pennsylvania School of Medicine
Department of Biostatistics and Epidemiology
622 Blockley Hall, 423 Guardian Drive
Philadelphia, PA 19104-6021

Abstract

In non-randomized studies, estimation of treatment effects generally requires adjustment for imbalances in observed covariates. One such method, based on the propensity score, is useful in many applications but may be biased when the assumption of strongly ignorable treatment assignment is violated. Because it is not possible to evaluate this assumption from the data, it is advisable to assess the sensitivity of conclusions to violations of strong ignorability. Lin et al [1] have implemented this idea by investigating how an unmeasured covariate may affect the conclusions of an observational study. We extend their method to assess sensitivity of the treatment hazard ratio to hidden bias under a range of covariate distributions. We derive simple formulas for approximating the true from the apparent treatment hazard ratio estimated under a specific survival model, and assess the validity of these formulas in simulation studies.

We demonstrate the method in an analysis of SEER-Medicare data on the effects of chemotherapy in elderly colon cancer patients.

Keywords: propensity score, accelerated failure time model, sensitivity analysis

* Corresponding author

1. INTRODUCTION

In randomized studies, there are, on average, no systematic differences in observed or unobserved covariates between treatment arms. In observational studies, where investigators do not control treatment assignment, the treated and control groups can differ in ways that significantly affect the outcomes under study. Large differences, if not properly controlled for, can severely bias estimated treatment effects.

Traditional methods for dealing with this problem include stratification, matching, and model-based covariate adjustment. Unfortunately, these methods all have their shortcomings: Matching fails when one cannot identify subjects who match on all important covariates; this can occur even when there are only a few such covariates. Stratification may fail when the number of strata is so large that not all strata contain both treated subjects and controls. Because the number of strata increases exponentially with the number of covariates, there is a risk of having such non-informative strata even when the number of covariates is quite small. Traditional model-based adjustment avoids these problems but depends on correct specification of the model relating the treatment and the covariates to the outcome. This can involve a perilous extrapolation if the distributions of the covariates in the comparison groups do not overlap substantially [2].

Rosenbaum and Rubin [3] introduced a new class of methods, based on the *propensity score*, for dealing with these problems. The propensity score is the conditional probability of assignment to a treatment given a vector of observed covariates. If one groups the subjects into strata based on their propensity scores, then treatment and control groups within the strata are balanced with respect to these observed potential confounders. Theoretical arguments show that subclassification into five propensity

score strata is usually adequate to remove over 90% of the bias [4]. Because exact adjustment using the propensity score will on average remove all of the bias, one can use propensity scores as a basis for matching, stratification, and regression adjustment. The method offers distinct advantages because it allows one to control for numerous covariates simultaneously by matching or stratifying on a single scalar variable [5]; this greatly simplifies model building and estimation.

Yet no adjustment method, even propensity scores, can completely eliminate potential bias from imbalance on covariates that do not appear in the data set. Hence, it is important to investigate how variations in assumptions about unmeasured confounders may affect study conclusions. This is the role of sensitivity analysis.

In observational studies we use sensitivity analysis to investigate how biases of various magnitudes affect inferences about treatment effects [6]. We posit a parameter, call it Γ, that measures the magnitude of the departure from a randomized experiment. For instance, suppose that a subject who is at level 1 of an unmeasured binary confounder has Γ times the probability of a subject who is at level 0 of being on the treatment arm. For $\Gamma = 1$, there is no hidden bias and the treatment assignment probabilities equal ½. For $\Gamma = 2$, matched subjects differ in their chances of receiving the treatment by a factor of 2, making one twice as likely as the other to receive the treatment. If small departures from $\Gamma = 1$ alter the conclusions of a study, then the study is said to be sensitive to hidden bias; if conclusions change only for large values of Γ, then results are insensitive. Note that sensitivity to small hidden biases does not imply that such biases are present, but that, if present, they can seriously alter conclusions.

Rosenbaum and Rubin [7] proposed methods for assessing the sensitivity to an unobserved binary covariate in observational studies where the outcome is binary. Lin et al. [1] extended these methods by deriving formulas to assess sensitivity to an unobserved normal covariate. Rosenbaum has also developed methods for analyzing sensitivity in matched case-control studies [8] and for matching with multiple controls [9].

Several studies in the 1980's investigated the effect of omitted covariates and misspecification of regression models in the context of survival analysis. Gail *et al*. [10] showed that when important covariates are omitted, certain nonlinear regression models lead to biased treatment effect estimates in randomized experiments. However, they found that for moderately censored data, when important covariates are omitted, treatment effect estimates from exponential survival models are less biased than those obtained from Cox proportional hazards models. Similarly, Chastang *et al*. [11] demonstrated that the effect of omitted covariates in survival models is modest unless the covariates are strongly prognostic. Solomon [12] investigated whether the choice of model family is critical in interpreting the importance of covariates in models of survival time and concluded that treatment effect estimates are approximately proportional under the Cox model and the accelerated failure time model so that qualitative inferences are robust to misspecification. Lagakos and Schoenfeld [13] investigated the effects of misspecifying a proportional hazards regression model on the partial likelihood score test when comparing two randomized treatments and found that omitting a balanced covariate has a negligible effect on the size of the score test but there is substantial reduction in power when the covariate effect is strong. Moreover, Struthers and Kalbfleisch [14] found that ignoring one of two independent covariates results in an attenuation of the

effect of the other covariate. Similarly, a study by Bretagnolle and Huber [15] supported the conclusions from previous studies by showing that in the Cox regression model, the effect of covariates included in the model are underestimated when other covariates are omitted. Subsequently, in the 1990's Hougard *et al*. (1994) and Keiding *et al.* [16] introduced a frailty model framework for describing the effect of omitted covariates. Their work suggested that accelerated failure time models may be better in accounting for heterogeneity in survival times due to unobserved or omitted covariates.

In this article, we extend the methods of Lin et al. [1] by deriving simple formulas for approximating the true from the apparent treatment hazard ratio estimated from a Weibull accelerated failure time model under a range of assumptions about the distribution of the omitted confounder. We use Monte-Carlo simulations to assess the accuracy of these approximations, and we apply our method to SEER-Medicare data to investigate the sensitivity of the treatment effect of chemotherapy on survival in elderly colon cancer patients.

2. STRONGLY IGNORABLE TREATMENT ASSIGNMENT

Causal inferences flow from models for the potential responses an individual could manifest under different treatments. Using the notation of Rosenbaum [17], let r_0 represent the response that one observes if the subject receives the control treatment, and r_1 be the response that one observes if the subject receives the experimental treatment. Because each subject receives only one treatment (z = 1 or 0), one observes r_1 for only those subjects who receive treatment (i.e., $z=1$), and r_0 for only those subjects who receive control. Although we can only conjecture about the causal effect of the treatment

for an individual, r_1-r_0, we can aspire to estimate the average causal effect, which is the difference between the *average* responses to treatment and control in a randomly selected subject:

$$E(r_1 - r_0)\,. \tag{2.1}$$

Treatment assignment is defined to be *strongly ignorable* if the response (r_0,r_1) is conditionally independent of the treatment assignment z given the observed covariates **x,** and the probability of receiving each treatment is positive:

$$\Pr(r_0, r_1, z \mid \mathbf{x}) = \Pr(r_0, r_1 \mid \mathbf{x}) \Pr(z \mid \mathbf{x}) \tag{2.2}$$

and

$$0 < \Pr(z = 1 \mid \mathbf{x}) < 1, \quad \forall \mathbf{x}\,. \tag{2.3}$$

For example, in the simplest randomized trial, subjects are assigned to treatment or control by the flip of a fair coin; that is $\Pr(z = 1 \mid \mathbf{x}) = 1/2$. Thus treatment assignment is strongly ignorable because the randomization probability is constant for all units regardless of the value of **x**. In more complicated randomized designs, randomization probabilities may depend on the values of observed covariates, but strong ignorability holds as long as these covariates are included in **x**.

When treatment assignment is strongly ignorable given **x**, we may obtain unbiased estimates of the average treatment effect, $E(r_1 - r_0)$, by matched sampling on **x**, subclassification on **x**, or covariance adjustment for **x**. First note that if a randomly selected treatment unit (z = 1) is compared to a randomly selected control unit (z = 0), the expected difference in response is

$$E(r_1 \mid z = 1) - E(r_0 \mid z = 0)\,, \tag{2.4}$$

which does not, in general, equal the average treatment effect of (2.1). Rosenbaum and Rubin [7] demonstrated that strongly ignorable treatment assignment can lead to unbiased estimates of the average treatment effect by considering the case of simple matched sampling. In this procedure, one randomly samples a value of **x** from the population and then samples a treatment unit and a control from all units with this value of **x**. The expected difference in outcomes is then

$$E_{\mathbf{x}}\{E(r_1 \mid z=1,\mathbf{x}) - E(r_0 \mid z=0,\mathbf{x})\} \tag{2.5}$$

where $E_{\mathbf{x}}\{\cdot\}$ is the expectation with respect to the distribution of **x** in the population. This does not generally equal the average treatment effect (2.1), but if we have strong ignorability, then (2.4) is equal to

$$E_{\mathbf{x}}\{E(r_1 \mid \mathbf{x}) - E(r_0 \mid \mathbf{x})\} \tag{2.6}$$

which does equal (2.1).

Rosenbaum and Rubin [7] have shown that if treatment assignment is strongly ignorable given **x**, then it is strongly ignorable given the propensity score, and moreover that in applications, where the exact propensity score is unknown, estimated propensity scores will serve as well. Although in observational studies we cannot directly test strong ignorability, by assessing the sensitivity of key inferences to departures from it we can determine the robustness and credibility of our conclusions.

3. CONFOUNDING IN THE ACCELERATED FAILURE TIME MODEL

The accelerated failure time (AFT) model provides an intuitive and flexible alternative to the Cox model in many applications [18]. In an AFT model, the covariates act multiplicatively on the time scale, implying that they affect the speed at which a

subject proceeds along the time axis [19]. When the predictor is a binary indicator of assignment to a test therapy, the survival time of a treated patient is thus a multiple of the survival time of an otherwise equivalent control, and so the effect of treatment is to either slow or accelerate the passage of time. Most simply, the model is

$$S_T(t) = S_C(\phi\, t)$$

where $S_T(t)$ and $S_C(t)$ are the survivor functions for subjects in the treated and control groups, respectively, and ϕ is the treatment effect. Thus values of ϕ less then 1 correspond to acceleration in the time to an event for the treated subject. The corresponding relationship between the hazard functions is

$$\lambda_T(t) = \phi\, \lambda_C(\phi\, t)\,.$$

Now let *Z* be the treatment indicator (1 if treated, 0 if not), **X** a set of measured covariates, *U* an unmeasured confounder, and *T* the survival time of interest. The AFT hazard function, conditional on (*Z,***X***,U*), is

$$\lambda(t \mid Z, \mathbf{X}, U) = \exp(\beta Z + \gamma_z U + \theta^{'}\mathbf{X}) \times \lambda_0\{\exp(\beta Z + \gamma_z U + \theta^{'}\mathbf{X})t\}$$

where $\lambda_0(\cdot)$ is a baseline hazard function and $(\beta, \gamma_0, \gamma_1, \theta)$ are regression parameters. Here γ_0 and γ_1 are the effects of *U* for the control and treated groups, respectively.

Similarly, the AFT hazard function, conditional on (*Z*,**X**) only, is

$$\lambda(t \mid Z, \mathbf{X}) = \exp(\beta^* Z + \theta^{*'}\mathbf{X}) \times \lambda^*{}_0\left\{\exp(\beta^* Z + \theta^{*'}\mathbf{X})t\right\}$$

where $\lambda_0^*(\cdot)$ is a baseline hazard function and (β^*, θ^*) are regression parameters. Lin et al. [1] call this the "reduced" model. β and β^* are the *true* and *apparent* exposure effects, respectively. Because we can estimate β^* from the observed data, our goal is to determine the relationship between β and β^*.

To completely determine an AFT model one needs to specify a probability distribution for the survival times. A common choice is the Weibull with scale parameter α and shape parameter ψ, whose baseline hazard function is

$$\lambda_0(t) = \alpha \psi t^{\psi - 1}.$$

Thus the hazard in a Weibull AFT model, conditional on (Z,$\mathbf{X}$,U) is

$$\lambda(t \mid Z, \mathbf{X}, U) = \alpha \psi t^{\psi - 1} \exp(\psi\beta Z + \psi\gamma_z U + \psi\theta' \mathbf{X}) \qquad (3.1)$$

and the hazard for the *reduced* AFT model, conditional on (Z,$\mathbf{X}$), is

$$\lambda(t \mid Z, \mathbf{X}) = \alpha \psi t^{\psi - 1} \exp(\psi\beta^* Z + \psi\theta^{*'} \mathbf{X}) \qquad (3.2)$$

In conventional notation, let $f(t \mid \cdot)$ be the conditional density function of T and $S(t \mid \cdot)$ be the conditional survival function of T, and let $F(u \mid Z, \mathbf{X})$ be the distribution function of U given Z and $\mathbf{X}$. The hazard function can be written as follows:

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

To determine the relationship between β and β^*, we derive the hazard given Z and $\mathbf{X}$ under the full model (3.1) directly from (3.3), then match terms with the apparent hazard under the reduced model, given in (3.2). Under (3.1), the numerator in (3.3) is

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

$$= \int_{-\infty}^{\infty} \alpha \ \psi \ t^{\psi-1} \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta' \mathbf{X}) \ \exp\left\{-\alpha \ t^{\psi} \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta' \mathbf{X})\right\} dF(u \mid Z, \mathrm{X})$$

The denominator is

$$\int_{-\infty}^{\infty} S(t \mid Z, \mathrm{X}, u) dF(u \mid Z, \mathrm{X})$$

$$= \int_{-\infty}^{\infty} \exp\left\{-\alpha \ t^{\psi} \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta' \mathbf{X})\right\} dF(u \mid Z, \mathrm{X})$$

Therefore, the hazard function in (3.3) can be rewritten as

$$\lambda(t \mid Z, \mathbf{X}) = \alpha \psi \, t^{\psi-1} \exp(\psi\beta Z + \psi\theta' \mathbf{X}) \, h(t; Z, \mathbf{X}) \tag{3.4}$$

where

$$h(t;Z,\mathrm{X}) = \frac{\int_{-\infty}^{\infty} \exp(\psi\gamma_z u)\, \exp\left\{-\alpha\, t^{\psi} \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta'\mathbf{X})\right\} dF(u \mid Z,\mathrm{X})}{\int_{-\infty}^{\infty} \exp\left\{-\alpha\, t^{\psi} \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta'\mathbf{X})\right\} dF(u \mid Z,\mathrm{X})}$$

Once one estimates ψ (the shape parameter of the Weibull distribution) and specifies the distribution of *U*, one can use (3.4) to determine an explicit formula for the hazard function. In section 3.1, we derive an explicit formula for the hazard function assuming that the unknown confounder follows a Gamma distribution. In sections 3.2, 3.3 and 3.4, we present similar formulas for normal, binomial and Poisson confounders. Derivations appear in the Appendix.

3.1 Unknown Gamma Confounder

Assume that, conditional on *Z* and **X**, the confounder *U* is Gamma with shape $\eta_{Z,\mathrm{X}}$ and scale $\tau_{Z,\mathbf{X}}$. We rewrite the hazard in (3.4) as follows:

$$\lambda(t \mid Z,\mathbf{X}) = \alpha\,\psi\, t^{\psi-1} \exp(\psi\beta Z + \psi\theta'\mathbf{X}) \frac{\int_{-\infty}^{\infty} e^{\gamma_z u} \exp\left\{-\Lambda_0(t) e^{\beta Z + \gamma_z u + \theta'\mathbf{X}}\right\} dF(u \mid Z,\mathbf{X})}{\int_{-\infty}^{\infty} \exp\left\{-\Lambda_0(t) e^{\beta Z + \gamma_z u + \theta'\mathbf{X}}\right\} dF(u \mid Z,\mathbf{X})}$$

$$= \alpha\,\psi\, t^{\psi-1}\exp(\psi\beta Z + \psi\theta^{'}\mathrm{X})$$
$$\times \frac{\int_{-\infty}^{\infty} e^{\psi\gamma_z u}\exp\left\{-\alpha\, t^{\psi}\exp(\psi\beta Z + \psi\gamma_z u + \psi\theta^{'}\mathbf{X})\right\}\frac{\tau_{z,\mathrm{x}}{}^{\eta_{z,\mathrm{x}}}}{\Gamma(\eta_{z,\mathrm{x}})}u^{\eta_{z,\mathrm{x}}-1}e^{-\tau_{z,\mathrm{x}}u}\,du}{\int_{-\infty}^{\infty}\exp\left\{-\alpha\, t^{\psi}\exp(\psi\beta Z + \psi\gamma_z u + \psi\theta^{'}\mathbf{X})\right\}\frac{\tau_{z,\mathrm{x}}{}^{\eta_{z,\mathrm{x}}}}{\Gamma(\eta_{z,\mathrm{x}})}u^{\eta_{z,\mathrm{x}}-1}e^{-\tau_{z,\mathrm{x}}u}\,du}$$

where $\Gamma(\eta)$ is the gamma function: $\Gamma(w) = \int_0^{\infty} y^{w-1}e^{-y}dy$.

If the events are rare then $\alpha\, t^{\psi}$ is small and the hazard simplifies to

$$\lambda(t \mid Z,\mathrm{X}) \approx \alpha\,\psi\, t^{\psi-1}\exp(\psi\beta Z + \psi\theta^{'}\mathrm{X})\int_{-\infty}^{\infty} e^{\psi\gamma_z u}\,\frac{\tau_{z,\mathrm{x}}{}^{\eta_{z,\mathrm{x}}}}{\Gamma(\eta_{z,\mathrm{x}})}u^{\eta_{z,\mathrm{x}}-1}e^{-\tau_{z,\mathrm{x}}u}\,du$$

$$= \alpha\,\psi\, t^{\psi-1}\exp(\psi\beta Z + \psi\theta^{'}\mathrm{X})\frac{\tau_{z,\mathrm{x}}{}^{\eta_{z,\mathrm{x}}}}{(\tau_{z,\mathrm{x}} - \psi\gamma_z)^{\eta_{z,\mathrm{x}}}}$$

Furthermore, if U is conditionally independent of $\mathbf{X}$ given Z, then under $\eta_{Z,\mathrm{X}} = \eta_Z$ and $\tau_{Z,\mathbf{X}} = \tau_Z$, the approximation becomes

$$\lambda(t \mid Z,\mathrm{X}) \approx \alpha\,\psi\, t^{\psi-1}\exp(\psi\beta Z + \psi\theta^{'}\mathrm{X})\frac{\tau_z{}^{\eta_z}}{(\tau_z - \psi\gamma_z)^{\eta_z}}$$

$$= \alpha\,\psi\, t^{\psi-1}\frac{\tau_0{}^{\eta_0}}{(\tau_0 - \psi\gamma_0)^{\eta_0}}\exp\left\{\left(\psi\beta + \log\left[\frac{\tau_1{}^{\eta_1}\big/(\tau_1 - \psi\gamma_1)^{\eta_1}}{\tau_0{}^{\eta_0}\big/(\tau_0 - \psi\gamma_0)^{\eta_0}}\right]\right)Z + \psi\theta'\mathrm{X}\right\}$$

Hence,

$$\beta \approx \beta^* - \frac{\log\left[\dfrac{\tau_1^{\eta_1} \big/ (\tau_1 - \psi\gamma_1)^{\eta_1}}{\tau_0^{\eta_0} \big/ (\tau_0 - \psi\gamma_0)^{\eta_0}}\right]}{\psi} \qquad (3.5)$$

In (3.5), γ_0 and γ_1 are the effects of U for the control and treated groups, respectively. Furthermore, η_0 and η_1 are the shape parameters of U, and τ_0 and τ_1 the scale parameters of U, for the control and treated groups, respectively. We can now estimate the true treatment effect β by first estimating the apparent treatment effect β^* from observed data and then applying (3.5). This formula also proves that when the sensitivity parameters are fixed, the variance of $\hat{\beta}$ is the same as that of $\hat{\beta}^*$, which allows us to easily calculate a confidence interval for β.

3.2 Unknown Normal Confounder

Now suppose that, conditional on Z and $\mathbf{X}$, the unknown confounder U is normally distributed with mean $\mu_{Z,\mathbf{X}}$ and unit variance. Then if the events are rare and U is conditionally independent of $\mathbf{X}$ given Z, under $\mu_{Z,\mathbf{X}} = \mu_Z$, the true treatment effect is

$$\beta \approx \beta^* - \left\{(\gamma_1\mu_1 - \gamma_0\mu_0) + \frac{\left(\psi\gamma^2{}_1 - \psi\gamma^2{}_0\right)}{2}\right\} \qquad (3.6)$$

Here γ_0 and γ_1 are the effects of U for the control and treated groups, respectively, and μ_0 and μ_1 are the means of U for the control and treated groups, respectively.

3.3 Unknown Binary Confounder

Assume that U is binary with success probability $\Pi_{Z,\mathbf{X}}$. If U is independent of $\mathbf{X}$ conditional on Z, which implies that $\Pi_{Z,\mathbf{X}}$ does not depend on $\mathbf{X}$, the treatment effect is

$$\beta \approx \beta^* - \frac{\log \dfrac{e^{\psi\gamma 1}\Pi_1 + (1-\Pi_1)}{e^{\psi\gamma 0}\Pi_0 + (1-\Pi_0)}}{\psi} \tag{3.7}$$

Here Π_1 and Π_0 are the prevalences of the unmeasured confounder in the treated and untreated groups, respectively, and γ_1 and γ_0 are the effects of the unmeasured binary confounder in the treated and untreated groups, respectively.

3.4 Unknown Poisson Confounder

Next suppose that, conditional on Z and $\mathbf{X}$, U is Poisson with mean $\kappa_{z,\mathbf{X}}$. As before, if events are rare and U is conditionally independent of $\mathbf{X}$ given Z so that $\kappa_{Z,\mathbf{X}} = \kappa_Z$, then

$$\beta \approx \beta^* - \frac{\kappa_1(e^{\psi\gamma 1} - 1) - \kappa_0(e^{\psi\gamma 0} - 1)}{\psi} \tag{3.8}$$

Here κ_0 and κ_1 are the means of U for the control and treated groups, respectively.

4. SIMULATIONS

In the previous section, we developed formulas to estimate β in the presence of an unmeasured confounder with a specified distribution (gamma, normal, binomial, or Poisson) assuming an AFT model. We now assess the accuracy of these approximations by Monte Carlo. We generated survival times under the AFT model from (3.1) with $\lambda_0(\cdot) = 1$, $\beta = 1$, $\theta = 1$, $\gamma_0 = \gamma_1$ and a Weibull distribution with scale $\alpha = 0.1$ and shape $\psi = 0.5$. We set 30% of the subjects to be in the treated group ($Z = 1$) and 70% in the control group ($Z = 0$) and assumed that **X** was normal with mean 1 for the treated group, mean 0 for the control group, and variance 1 for both groups. We generated censoring times from the Uniform (0,ω) distribution where ω was chosen to fix the censoring rate. Results corresponding to the four models appear in tables 1 through 4. In each simulation, there were 1000 samples, each containing 1000 subjects. For each sample, we calculated the maximum partial likelihood estimate $\hat{\beta}^*$ under the reduced model and then replaced β^* in the appropriate formulas by $\hat{\beta}^*$ to estimate $\hat{\beta}$.

With binary U, our results suggest that (3.7) is a good approximation under a wide range of censoring rates. When U is normal, however, the approximation (3.6) is adequate only when the censoring rate is over 90% or when the effect of the unmeasured confounder on survival time is small ($\gamma < 0.75$). When U is Poisson, the approximation (3.8) is satisfactory when the censoring percentage is over 75%. Finally, when U has a gamma distribution, (3.5) is a good approximation when the censoring rate is over 50%

or when the effect of U on T is small ($\gamma < 0.25$). Overall, if the censoring rate is high or the effect of U on T is small, then the approximations are sufficiently unbiased for practical purposes.

5. COLON CANCER STUDY

The aim was to estimate the effect of adjuvant chemotherapy with 5-flourouracil (5FU) in extending survival in older, node-positive colon cancer patients. Our data source was the Surveillance Epidemiology and End Results - Medicare (SEER-Medicare) database that links cancer registry information with Medicare claims data [20]. SEER collects information on tumor location, stage of disease, and demographics (e.g. age, race, sex, and area of residence), together with primary surgical and radiation treatment and survival; it covers roughly 14% of the US population. The linked Medicare files contain extensive diagnostic, treatment, and cost data.

We identified all individuals in the database who had a first primary diagnosis of colon cancer between 1992 and 1996. Other selection criteria included: age $\geq$ 65 years; diagnosis of American Joint Cancer Committee (AJCC) stage III colon cancer; no HMO enrollment the year before diagnosis (the Health Care Finance Administration does not collect claims data from risk-based HMOs) and coverage with both Medicare Parts A and B in the year before diagnosis; survival for 120 days beyond the date of diagnosis; and confirmed positive regional nodes on surgical resection of the tumor.

We measured the presence of other illnesses by the Deyo adaptation of the Charlson comorbidity index. For those who died by the end of follow-up (April 15,

1999), survival was the interval from the cancer diagnosis date to the Medicare date of death. We classified those surviving beyond April 15, 1999 as censored.

We hypothesized that 5FU treatment would be associated with longer survival, but we were concerned that the subjects receiving 5FU would be quite different from the others in terms of covariates that are likely to predict survival. Of a total of 4,768 patients in the dataset, 2,479 received 5FU and 2,289 did not. An initial analysis showed significant differences in several background covariates (Table 5). Note, for example, that younger, urban and white patients were all more likely to be treated.

We estimated propensity scores from a logistic regression including 22 potential predictors. A patient's propensity score is his estimated probability of being treated with 5FU given his background covariates. We assigned patients to five strata based on quintiles of the distribution of the propensity scores. Patients in the first quintile have the lowest propensity scores, indicating that they were the types of patients who were less likely to receive 5FU (although some did). Those assigned to each successive quintile were more likely to receive 5FU (although in each case some did not).

We compared treated and control groups on each of the 22 covariates using a two-way ANOVA with propensity quintile and treatment group (5FU/no 5FU) as the factors. As expected, the two groups were quite similar within quintiles.

We estimated the effect of treatment on survival separately within each quintile (Table 6). For instance, in quintile 1, the hazard of death for those who received 5FU is about 81% of the hazard for those who did not receive 5FU. Each of the five hazard ratios is less than 1, suggesting that 5FU improves survival. All ratios are significant except for quintile 1, which is the stratum least likely to receive 5FU.

Next we estimated the treatment effect by an AFT model that included the propensity score quintile (to account for the initial imbalance in the measured covariates) and age at diagnosis, which was not perfectly balanced even after propensity score adjustment. Our results show that the hazard of death for those who received 5FU is about 66% of the hazard for those who did not.

Although the propensity score adjustment included 22 covariates, which presumably captured most of the potential confounding, the possibility still exists that the apparent association between treatment and response resulted from some unknown or unmeasured covariates that were associated with both treatment and outcome. An example of a binary unmeasured confounder in the 5FU study is performance status or functional impairment. Performance status has been linked to survival in cancer patients and is likely to be associated with the decision to treat with chemotherapy (i.e., the better the status, the more likely a patient is to receive chemotherapy). Thus it is possible that decreased performance status is unequally distributed between the treatments, such that its prevalence was higher in the group not treated with 5FU, and that it is associated with an increased hazard of death. This could bias even the propensity-score adjusted analysis. Our methodology allows us to model this in a formal sensitivity analysis.

We estimated the hazard ratio for 5FU treatment to be 0.66 with a 95% confidence interval of (0.60, 0.73). We assumed in our sensitivity analysis that the unmeasured confounder is more prevalent in the group not treated with 5FU and that it is associated with decreased survival, which would bias results toward the null. Important parameters in the sensitivity analysis include the prevalence of the unknown confounder

in the treated and control groups in addition to the relative hazard of death associated with the unmeasured confounder.

Our analyses (Table 7) show that a binary confounder would need to increase the risk of death by at least 50% (hazard of at least 1.5) before it would alter our results. Moreover, a confounder with a hazard ratio of 1.5 or greater would have to be significantly unbalanced between groups to alter results. For example, a confounder with a hazard of 1.75 would need to be more than 5 times as prevalent in the control group to render the treatment effect non-significant. Thus the beneficial effect of 5FU appears to be robust to departures from strong ignorability.

6. DISCUSSION

We have demonstrated a methodology for assessing sensitivity of inferences to departures from strong ignorability of treatment assignment, within the context of the accelerated failure time model for survival. The adjustments are simple to compute, in that they merely require one to fit the AFT model without the confounder, then specify the parameters of the confounding process and do a simple linear adjustment. The range of validity of the approximations depends on the model for the confounder, but all are valid as long as the fraction of events is small and the effect of the confounder is modest. The latter is a less important issue because we are generally more interested in precise quantitation of sensitivity to small departures from the strongly ignorable model. The assumption of rare events in the proposed model is not a serious limitation because one of the main reasons that observational studies are conducted is precisely because the event of interest is rare. Moreover, our simulations suggest that in the case of unmeasured

binary confounders, the approximations are adequate even when the fraction of events is not small. When the approximation is suspect, one can develop adjustments by straightforward simulation approaches. Although we have applied our method in the context of a propensity score analysis, the adjustment formulas apply equally well to an AFT model estimated in the context of a matched, stratified or purely model-based analysis.

Our example demonstrates a practical application of the method in a study of survival of older colon cancer patients. Our analysis posited that an underlying binary performance status variable, which was not explicitly coded in the dataset, could have influenced the decision to administer adjuvant chemotherapy and been predictive of outcome. In other applications, where such a rich array of covariates was not available, one might hypothesize that the candidate confounder was some count variable (modeled with the Poisson), a symmetric continuous variable (modeled with the normal), or a skewed continuous variable (modeled with the gamma). We have derived adjustment formulas for each of these situations.

APPENDIX

Unknown Binary Confounder

Assume that the unknown confounder U is binary with success probability $\Pi_{Z,\mathbf{X}}$.

Then

$$h(t;Z,\mathbf{X}) = \frac{e^{\psi\gamma z}\exp\left\{-\alpha t^{\psi}\exp(\psi\beta Z+\psi\gamma_z+\psi\theta'\mathbf{X})\right\}\Pi_{z,\mathbf{x}}+\exp\left\{-\alpha t^{\psi}\exp(\psi\beta Z+\psi\theta'\mathbf{X})\right\}(1-\Pi_{z,\mathbf{x}})}{\exp\left\{-\alpha t^{\psi}\exp(\psi\beta Z+\psi\gamma_z+\psi\theta'\mathbf{X})\right\}\Pi_{z,\mathbf{x}}+\exp\left\{-\alpha t^{\psi}\exp(\psi\beta Z+\psi\theta'\mathbf{X})\right\}(1-\Pi_{z,\mathbf{x}})}$$

If events are rare, implying that αt^{ψ} is small, the approximate hazard is

$$\lambda(t\mid Z,\mathbf{X}) \approx \alpha\psi t^{\psi-1}\exp(\psi\beta Z+\psi\theta'\mathbf{X})\{e^{\psi\gamma z}\Pi_{Z,\mathbf{X}}+(1-\Pi_{Z,\mathbf{X}})\}$$

For simplicity, assume that U is independent of $\mathbf{X}$ conditional on Z, which implies that $\Pi_{Z,\mathbf{X}}$ does not depend on $\mathbf{X}$. Thus letting $\Pi_{Z,\mathbf{X}}=\Pi_Z$ ($Z=0$ or 1),

$$\lambda(t\mid Z,\mathbf{X}) \approx \alpha\psi t^{\psi-1}\exp(\psi\beta Z+\psi\theta'\mathbf{X})\{e^{\psi\gamma z}\Pi_Z+(1-\Pi_Z)\}$$

$$=\alpha\psi t^{\psi-1}\exp(\psi\beta Z+\psi\theta'\mathbf{X})\left\{e^{\psi\gamma 0}\Pi_0+(1-\Pi_0)\right\}\times\exp\left\{Z\log\left[\frac{e^{\psi\gamma 1}\Pi_1+(1-\Pi_1)}{e^{\psi\gamma 0}\Pi_0+(1-\Pi_0)}\right]\right\}$$

$$=\alpha\psi t^{\psi-1}\left\{e^{\psi\gamma 0}\Pi_0+(1-\Pi_0)\right\}\times\exp\left(\left\{\psi\beta+\log\left[\frac{e^{\psi\gamma 1}\Pi_1+(1-\Pi_1)}{e^{\psi\gamma 0}\Pi_0+(1-\Pi_0)}\right]\right\}Z+\psi\theta'\mathbf{X}\right)$$

where $\beta^* = \dfrac{\psi\beta + \log \dfrac{e^{\psi\gamma 1}\Pi_1 + (1-\Pi_1)}{e^{\psi\gamma 0}\Pi_0 + (1-\Pi_0)}}{\psi}$.

Hence,

$$\beta \approx \beta^* - \frac{\log \dfrac{e^{\psi\gamma 1}\Pi_1 + (1-\Pi_1)}{e^{\psi\gamma 0}\Pi_0 + (1-\Pi_0)}}{\psi}$$

Unknown Normal Confounder

Suppose that, conditional on Z and $\mathbf{X}$, U is normal with mean $\mu_{Z,\mathbf{X}}$ and unit variance. Then

$$\lambda(t \mid Z, \mathbf{X}) = \alpha\, \psi\, t^{\psi-1} \exp(\psi\beta Z + \psi\theta^{'}\mathbf{X}) \times h(t; Z, \mathbf{X})$$

where

$$h(t; Z, \mathbf{X}) = \frac{\displaystyle\int_{-\infty}^{\infty} \exp(\psi\gamma_z u)\, \exp\left\{-\alpha\, t^{\psi}\, \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta^{'}\mathbf{X})\right\} dF(u \mid Z, \mathbf{X})}{\displaystyle\int_{-\infty}^{\infty} \exp\left\{-\alpha\, t^{\psi}\, \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta^{'}\mathbf{X})\right\} dF(u \mid Z, \mathbf{X})}$$

The numerator of $h(t; Z, \mathbf{X})$ simplifies to

$$\int_{-\infty}^{\infty} \exp(\psi\gamma_z u)\ \exp\left\{-\alpha\, t^{\psi}\ \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta^{'}\mathbf{X})\right\}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(u-\mu_{Z,\mathbf{X}})^2}{2}\right\}du$$

$$= \exp(\psi\gamma_z\mu_{Z,\mathbf{X}})\exp\frac{(\psi\gamma_z)^2}{2}\int_{-\infty}^{\infty}\exp\left\{-\alpha\, t^{\psi}\ \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta^{'}\mathbf{X})\right\}$$
$$\times\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(u-\mu_{Z,\mathbf{X}}-\psi\gamma_z)^2}{2}\right\}du$$

Again, if events are rare, $\alpha\, t^{\psi}$ is small and $h(t;Z,\mathbf{X})$ simplifies to

$$h(t;Z,\mathbf{X}) \approx \exp\left\{\psi\gamma_z\mu_{Z,\mathbf{X}} + \frac{(\psi\gamma_z)^2}{2}\right\}$$

Thus,

$$\lambda(t \mid Z,\mathbf{X}) \approx \alpha\,\psi\, t^{\psi-1}\exp(\psi\beta Z + \psi\theta^{'}\mathbf{X})\times\exp\left\{\psi\gamma_z\mu_{Z,\mathbf{X}} + \frac{(\psi\gamma_z)^2}{2}\right\}$$

Furthermore, if U is conditionally independent of $\mathbf{X}$ given Z, then under $\mu_{Z,\mathbf{X}} = \mu_Z$,

$$\lambda(t \mid Z,\mathbf{X}) \approx \alpha\,\psi\, t^{\psi-1}\exp\left\{\psi\beta Z + \psi\theta^{'}\mathbf{X} + \psi\gamma_z\mu_Z + \frac{(\psi\gamma_z)^2}{2}\right\}$$

$$= \alpha \psi t^{\psi-1} \exp\left(\psi\gamma_0\mu_0 + \frac{\psi^2 {\gamma_0}^2}{2} \right)$$
$$\times \exp\left(\left[\psi\beta + \psi\gamma_1\mu_1 - \psi\gamma_0\mu_0 + \frac{\psi^2 {\gamma_1}^2}{2} - \frac{\psi^2 \gamma^2 0}{2} \right] Z + \psi\theta' \mathbf{X} \right)$$

where

$$\beta^* \approx \frac{\psi\beta + \psi\gamma_1\mu_1 - \psi\gamma_0\mu_0 + \frac{\psi^2 {\gamma_1}^2}{2} - \frac{\psi^2 \gamma^2 0}{2}}{\psi}$$

Hence

$$\beta \approx \beta^* - \left\{ \left(\gamma_1\mu_1 - \gamma_0\mu_0\right) + \frac{\left(\psi\gamma^2 1 - \psi\gamma^2 0\right)}{2} \right\}$$

Unknown Poisson Confounder

Suppose that, conditional on Z and $\mathbf{X}$, U is Poisson with mean $\kappa_{z,\mathbf{x}}$. We write the hazard as

$$\lambda(t \mid Z, \mathbf{X}) = \alpha \psi t^{\psi-1} \exp(\psi\beta Z + \psi\theta' \mathbf{X}) \times h(t; Z, \mathbf{X})$$

where

$$h(t;Z,\mathbf{X}) = \frac{\sum_{u=0}^{\infty} e^{\psi\gamma_z u} \exp\left\{-\alpha\, t^{\psi} \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta'\mathbf{X})\right\} e^{-\kappa_{z,\mathbf{x}}} \frac{\kappa_{z,\mathbf{x}}{}^{u}}{u!}}{\sum_{u=0}^{\infty} \exp\left\{-\alpha\, t^{\psi} \exp(\psi\beta Z + \psi\gamma_z u + \psi\theta'\mathbf{X})\right\} e^{-\kappa_{z,\mathbf{x}}} \frac{\kappa_{z,\mathbf{x}}{}^{u}}{u!}}$$

As before, if events are rare, then $\alpha\, t^{\psi}$ is small and the formula for the hazard will simplify. We also assume that U is conditionally independent of $\mathbf{X}$ given Z so that $\kappa_{Z,\mathbf{X}} = \kappa_Z$. Under these assumptions, the hazard function is approximately

$$\lambda(t \mid Z,\mathbf{X}) \approx \alpha\,\psi\, t^{\psi-1} \exp(\psi\beta Z + \psi\theta'\mathbf{X}) \sum_{u=0}^{\infty} e^{\psi\gamma_z u} e^{-\kappa_z} \frac{\kappa_z{}^{u}}{u!}$$

$$= \alpha\,\psi\, t^{\psi-1} \exp(\psi\beta Z + \psi\theta'\mathbf{X}) e^{-\kappa_z} \sum_{u=0}^{\infty} \frac{\left(\kappa_z e^{\psi\gamma_z}\right)^{u}}{u!}$$

$$= \alpha\,\psi\, t^{\psi-1} \exp(\psi\beta Z + \psi\theta'\mathbf{X}) e^{\kappa_z (e^{\psi\gamma_z} - 1)}$$

$$= \alpha\,\psi\, t^{\psi-1} \exp(\kappa_0 e^{\psi\gamma 0} - \kappa_0) \exp\{\left(\psi\beta + \kappa_1 e^{\psi\gamma 1} - \kappa_0 e^{\psi\gamma 0} - \kappa_1 + \kappa_0\right) Z + \psi\theta'\mathbf{X}\}$$

Hence

$$\beta^* \approx \beta + \frac{\kappa_1 e^{\psi\gamma 1} - \kappa_0 e^{\psi\gamma 0} - \kappa_1 + \kappa_0}{\psi}$$

and

$$\beta \approx \beta^* - \frac{\kappa_1 (e^{\psi\gamma 1} - 1) - \kappa_0 (e^{\psi\gamma 0} - 1)}{\psi}$$

ACKNOWLEDGMENTS

This study used the linked SEER-Medicare Database. The interpretation and reporting of these data are the sole responsibility of the authors, who gratefully acknowledge the efforts of the Applied Research Program, NCI; the Office of Information Services, and the Office of Strategic Planning, CMS; Information Management Services (IMS), Inc.; and the Surveillance, Epidemiology, and End Results (SEER) Program tumor registries. The USPHS supported this research under grants HL-68074 and CA09529. We thank Drs. Vijaya Sundararajan and Al Neugut for assistance with the data and for many stimulating discussions.

REFERENCES

1. Lin DY, Psaty BM, Kronmal RA. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics* 1998;**54**(3):948-63.

2. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998;**17**(19):2265-81.

3. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;**70**:41-55.

4. Cochran WG. The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics* 1968;**24**(2):295-313.

5. Rubin DB, Thomas N. Matching using estimated propensity scores: relating theory to practice. *Biometrics* 1996;**52**(1):249-64.

6. Rosenbaum PR. Discussing hidden bias in observational studies. *Ann Intern Med* 1991;**115**(11):901-5.

7. Rosenbaum PR, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society, Series B* 1983b;**45**:212-8.

8. Rosenbaum PR. Sensitivity analysis for matched case-control studies. *Biometrics* 1991;**47**(1):87-100.

9. Rosenbaum PR. Sensitivity analysis for matching with multiple controls. *Biometrika* 1988;**75**:577-81.

10. Gail MH, Wieand S, Piantadosi S. Biased estimates of treatment effects in randomized experiments with non-linear regressions and omitted covariates. *Biometrika* 1984;**71**:431-444.

11. Chastang C, Byar D, Piantadosi S. A quantitative study of the bias in estimating the treatment effect caused by omitting a balanced covariate in survival models. *Stat Med* 1988;**7**(12):1243-55.

12. Solomon PJ. Effect of misspecification of regression models in the analysis of survival data. *Biometrika* 1986;**71**:291-298.

13. Lagakos SW, Schoenfeld DA. Properties of proportional-hazards score tests under misspecified regression models. *Biometrics* 1984;**40**(4):1037-48.

14. Struthers CA, Kalbfleisch JD. Misspecified proportional hazard models. *Biometrika* 1986;**73**:363-369.

15. Bretagnolle J, Huber-Carol C. Effects of omitting covariates in Cox's model for survival data. *Scandinavian Journal of Statistics* 1998;**15**:125-138.

16. Hougaard P, Myglegaard P, Borch-Johnsen K. Heterogeneity models of disease susceptibility, with application to diabetic nephropathy. *Biometrics* 1994;**50**(4):1178-88.

17. Rosenbaum PR. From association to causation in observational studies: the role of tests of strongly ignorable treatment assignment. *J Am Stat Assoc* 1984;**79**:41-8.

18.Wei LJ. The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis. *Stat Med* 1992;**11**(14-15):1871-9.

19. Collet D. *Modelling Survival Data in Medical Research*. London: Chapman & Hall, 1994.

20. Sundararajan V, Mitra N, Jacobson JS*, et al.* Survival associated with 5-fluorouracil-based adjuvant chemotherapy among elderly patients with node-positive colon cancer. *Ann Intern Med* 2002;**136**(5):349-57.

Table 1

Monte Carlo estimates for the sampling means of $\hat{\beta}$ and the coverage probabilities for the 95% intervals (two-sided) for β.

Censoring	γ	$\hat{\beta}$ (coverage probability)			
		Binary $\pi 0 = 0.3, \pi 1 = 0.8$	Normal $\mu_0 = 0, \mu_1 = 1, \sigma^2 = 1$	Poisson $\kappa_0 = 2, \kappa_1 = 4$.	Gamma $\eta_0 = 2$ $\eta_1 = 5,$ $\tau_0 = 1.5$ $\tau_1 = .25$
90%	0.25	1.00	0.98	0.99	1.00
90%	0.50	1.01	0.97	0.97	1.03
90%	0.75	1.01	0.91	0.97	1.03
90%	1.00	1.03	0.85	0.96	1.08
75%	0.25	1.01	0.90	0.96	1.06
75%	0.50	1.01	0.86	0.92	1.06
75%	0.75	1.02	0.81	0.89	1.10
75%	1.00	1.02	0.76	0.88	1.14
50%	0.25	0.99	0.85	0.85	1.14
50%	0.50	0.99	0.80	0.81	1.15
50%	0.75	0.98	0.74	0.79	1.16
50%	1.00	0.97	0.72	0.73	1.20

Table 2

***Normal Confounder*. Monte Carlo estimates for the sampling means of $\hat{\beta}$ and the 95% confidence intervals.** $\mu_0 = 0, \mu_1 = 1, \sigma^2 = 1$

Censoring	γ	$\hat{\beta}$	95% CI
90%	0.25	0.98 (0.86, 1.10)	
90%	0.50	0.97(0.83, 1.11)	
90%	0.75	0.91(0.78, 1.04)	
90%	1.00	0.85(0.70, 1.00)	
75%	0.25	0.90(0.80, 1.01)	
75%	0.50	0.86(0.70, 1.02)	
75%	0.75	0.81(0.69, 0.93)	
75%	1.00	0.76(0.62, 0.90)	
50%	0.25	0.85(0.73, 0.97)	
50%	0.50	0.80(0.69, 0.91)	
50%	0.75	0.74(0.63, 0.85)	
50%	1.00	0.72(0.59, 0.85)	

Table 3

***Poisson Confounder*. Monte Carlo estimates for the sampling means of $\hat{\beta}$ and the 95% confidence intervals. Assume $\kappa_0 = 2$ and $\kappa_1 = 4$.**

Censoring	γ	$\hat{\beta}$	95% CI
90%	0.25	0.99	(0.87, 1.11)
90%	0.50	0.97	(0.84, 1.10)
90%	0.75	0.97	(0.86, 1.08)
90%	1.00	0.96	(0.82, 1.10)
75%	0.25	0.96	(0.80, 1.12)
75%	0.50	0.92	(0.77, 1.07)
75%	0.75	0.89	(0.73, 1.05)
75%	1.00	0.88	(0.73,1.03)
50%	0.25	0.85	(0.72, 0.98)
50%	0.50	0.81	(0.68, 0.94)
50%	0.75	0.79	(0.65, 0.93)
50%	1.00	0.73	(0.59, 0.87)

Table 4

***Gamma Confounder*. Monte Carlo estimates for the sampling means of $\hat{\beta}$ and the 95% confidence intervals. Assume $\eta_0 = 2$, $\eta_1 = 5$, and $\tau_0 = 1/2$ and $\tau_1 = 1/4$.**

Censoring	γ	$\hat{\beta}$	95% CI
90%	0.25	1.00	(0.88, 1.12)
90%	0.50	1.03	(0.90, 1.16)
90%	0.75	1.03	(0.92, 1.14)
90%	1.00	1.08	(0.94, 1.22)
75%	0.25	1.06	(0.90, 1.22)
75%	0.50	1.06	(0.91, 1.21)
75%	0.75	1.10	(0.94, 1.26)
75%	1.00	1.14	(0.99, 1.29)
50%	0.25	1.14	(1.01, 1.27)
50%	0.50	1.15	(1.02, 1.28)
50%	0.75	1.16	(1.02, 1.30)
50%	1.00	1.20	(1.06, 1.34)

Table 2
Comparison of background covariates between treated and untreated groups:

	5-FU (N=2479) mean (sd)	No 5-FU (N=2289) mean (sd)	95% CI for Difference in Means
Age at Diagnosis	73.90 (5.10)	78.17 (7.20)	(3.88, 4.66)*
Gender (Male)	0.47 (0.50)	0.41 (0.49)	(-0.10, -0.04)*
Race			
White	0.86 (0.34)	0.82 (0.38)	(-0.06, -0.02)*
Black	0.06 (0.24)	0.08 (0.28)	(0.01, 0.04)*
Hispanic	0.04 (0.19)	0.03 (0.18)	(-0.02, 0.01)
Other	0.04 (0.19)	0.06 (0.24)	(0.01, 0.04)*
Urban Residence	0.87 (0.33)	0.81 (0.40)	(-0.10, -0.05)*
Year of Diagnosis	93.99 (1.43)	93.94 (1.41)	(-0.13, 0.04)
Lymph Nodes			
1-3 nodes	0.66 (0.47)	0.69 (0.46)	(0.01, 0.06)*
4+ nodes	0.31 (0.45)	0.27 (0.46)	(-0.06, -0.01)*
unknown	0.03 (0.18)	0.03 (0.18)	(-0.01, 0.01)
Deyo-Charlson Comorbidity (any/none)	0.33 (0.67)	0.47 (0.86)	(0.09, 0.18)*

* Significant at 0.05 level

Table 3
Estimates of hazard ratios within quintiles

Quintile	Percentage of patients treated	Mean age (range)	Hazard Ratio	95% confidence interval
1	11%	87 (73 -101)	0.81	(0.62, 1.05)
2	38%	80 (67 – 88)	0.55	(0.46, 0.66)
3	59%	76 (66 – 83)	0.64	(0.54, 0.77)
4	74%	73 (65 – 79)	0.75	(0.61, 0.92)
5	83%	69 (65 – 74)	0.71	(0.55, 0.91)
Overall	52%	76 (65 – 103)	0.66	(0.60, 0.73)

Table 4
Assessing the sensitivity of the hazard ratio* to an unmeasured binary confounder:

Prevalence of unmeasured confounder in control group	Prevalence of unmeasured confounder in 5FU group	Hazard ratio of unmeasured confounder	Hazard ratio after adjustment for unmeasured confounder	95% Confidence Interval for Hazard ratio
0.9	0.1	3	1.55	(1.41, 1.69)
0.8	0.1	3	1.44	(1.31, 1.57)
0.7	0.1	3	1.33	(1.21, 1.45)
0.6	0.1	3	1.22	(1.11, 1.33)
0.5	0.1	3	1.11	(1.00, 1.21)
0.9	0.1	2	1.15	(1.05, 1.25)
0.8	0.1	2	1.08	(0.99, 1.19)
0.7	0.1	2	1.02	(0.94, 1.12)
0.6	0.1	2	0.96	(0.88, 1.06)
0.5	0.1	2	0.90	(0.83, 0.99)
0.9	0.5	2	0.84	(0.77, 0.92)
0.9	0.1	1.75	1.03	(0.94, 1.13)
0.9	0.5	1.75	0.81	(0.74, 0.88)
0.5	0.1	1.75	0.85	(0.77, 0.93)
0.9	0.1	1.5	0.92	(0.84, 1.00)
0.9	0.5	1.5	0.77	(0.70, 0.84)
0.5	0.1	1.5	0.79	(0.72, 0.86)
0.9	0.1	1.25	0.79	(0.72, 0.87)
0.9	0.5	1.25	0.72	(0.66, 0.79)
0.5	0.1	1.25	0.73	(0.66, 0.80)
0.9	0.1	1.1	0.72	(0.65, 0.78)
0.9	0.5	1.1	0.69	(0.63, 0.75)
0.5	0.1	1.1	0.69	(0.63, 0.75)

* Hazard ratio before adjusting for confounder = 0.66, 95%CI:(0.60, 0.73)