Extension of a Cox proportional hazards cure model when cure information is partially known

YU WU
K&L Consulting Services, Inc., Fort Washington, PA 19034, USA

YONG LIN*, SHOU-EN LU
Department of Biostatistics, School of Public Health, Rutgers, The State University of New Jersey, 683 Hoes Lane West, Piscataway, NJ 08854, USA
Biometrics Division, Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, 120 Albany Street, New Brunswick, NJ 08901, USA
linyo@sph.rutgers.edu

CHIN-SHANG LI
Department of Public Health Sciences, Division of Biostatistics, University of California, Davis, CA 95616, USA

WEICHUNG JOE SHIH
Department of Biostatistics, School of Public Health, Rutgers, The State University of New Jersey, 683 Hoes Lane West, Piscataway, NJ 08854, USA
Biometrics Division, Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, 120 Albany Street, New Brunswick, NJ 08901, USA

SUMMARY

When there is evidence of long-term survivors, cure models are often used to model the survival curve. A cure model is a mixture model consisting of a cured fraction and an uncured fraction. Traditional cure models assume that the cured or uncured status in the censored set cannot be distinguished. But in many practices, some diagnostic procedures may provide partial information about the cured or uncured status relative to certain sensitivity and specificity. The traditional cure model does not take advantage of this additional information. Motivated by a clinical study on bone injury in pediatric patients, we propose a novel extension of a traditional Cox proportional hazards (PH) cure model that incorporates the additional information about the cured status. This extension can be applied when the latency part of the cure model is modeled by the Cox PH model. Extensive simulations demonstrated that the proposed extension provides more efficient and less biased estimations, and the higher efficiency and smaller bias is associated with higher sensitivity and specificity of diagnostic procedures. When the proposed extended Cox PH cure model was applied to the motivating example, there was a substantial improvement in the estimation.

Keywords: Cure model; Expectation-maximization (EM) algorithm; Proportional hazards; Relative efficiency; Sensitivity and specificity.

*To whom correspondence should be addressed.

© The Author 2014. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.
1. Introduction

A common assumption in survival analysis is that if there is no censoring, an event of interest will eventually occur. Under this assumption the survival curve will eventually decrease to 0. There are, however, many examples where this assumption does not hold. In clinical settings, Withers and others (1995) showed Kaplan–Meier (K–M) curves for local control by external beam radiation therapy for squamous carcinoma of tonsil as a function of T- and N-stage. Patients were followed up for a minimum of 5 years. For each of the K–M curves, a plateau was reached in about 1–3 years of follow-up. Taylor (1995) pointed out that because of the kinetics of tumor growth, it is unlikely to have any recurrence later than 5 years after radiation treatment, although a few of them might still have recurrences (related to the sensitivity and specificity of a 5-year criterion). In sociology, Yamaguchi (1992) illustrated several examples of “cured” subjects: marriage, divorce, birth of second child, and career shift. Using career shift as an example, after an individual has been in an industry for more than 10 years, she/he can be considered “cured” because it is unlikely for the individual to shift in career to another industry, although a few might still shift their careers (related to the sensitivity and specificity of a 10-year criterion). In general, when an event of interest is not death, the subjects whose event of interest will not occur are termed cured subjects in the literature.

The Cox proportional hazards (PH) model (Cox, 1972) has been widely used for survival analysis. When these models are specified parametrically, the underlying assumption is that the event of interest will eventually occur. This assumption is not appropriate for cured subjects. The Cox PH model can be used potentially for the cure information by setting the survival function to 0 after a time threshold. But, when modeling this way, long-term survivors cannot be distinguished from cured subjects.

Alternative approaches using cure models have been studied by many authors. We denote $T$ as a non-negative random variable for the failure time, $\mathbf{x}$ and $\mathbf{z}$ as the covariate vectors, $\pi(\mathbf{z})$ as the uncured probability for a subject, and $S(t|\mathbf{x}, \mathbf{z})$ as the survival function for $T$, respectively. Let $f_u(t|\mathbf{x})$ and $S_u(t|\mathbf{x})$ be the probability density function (pdf) and the survival function for uncured subjects. The Cox PH cure model can be written as a mixture model in terms of the survival function:

$$S(t|\mathbf{x}, \mathbf{z}) = \pi(\mathbf{z})S_u(t|\mathbf{x}) + [1 - \pi(\mathbf{z})].$$

(1.1)

It is noted that $f(t|\mathbf{x}, \mathbf{z}) = -dS(t|\mathbf{x}, \mathbf{z})/dt = \pi(\mathbf{z})f_u(t|\mathbf{x})$. Here, the “incidence” part $\pi(\mathbf{z})$ is modeled by logistic regression. The “latency” part $f_u(t|\mathbf{x})$ or $S_u(t|\mathbf{x})$ is modeled by the Cox PH model (Kuk and Chen, 1992; Peng and Dear, 2000; Sy and Taylor, 2000). Parametric cure models can achieve the greatest efficiency in estimation if their distributional assumptions are satisfied. However, verifying these assumptions can be challenging in practice. Semi-parametric models do not require a distributional assumption, but may lose efficiency in estimation compared with a parametric model when a distribution can be correctly identified.

All the cure modeling to date assumes that cured and uncured subjects cannot be distinguished in the censored subset. But in some cases, we are able to get the “cured” status of a subject established either fully or partially. For example, in the radiology example mentioned above, if subjects are unlikely to have events when a recurrence has not happened in 5 years of follow-up, we can be almost sure that subjects who have been followed up more than 5 years without a local recurrence event are cured subjects. Our motivating example is a study on bone injury in pediatric patients (Leary and others, 2009). Medical charts of 157 children were reviewed to identify the incidence of premature physeal closure (PPC) following physeal fractures of the distal end of the tibia. For 16 children, their PPCs were identified. Other children were censored. Among the censored children, closure of the growth plate was noted without the events of PPC, thus they could be viewed as cured. Because the K–M curve for PPC in Figure 1(a) shows that there is a clear cure indication in this dataset, it is appropriate to use a cure model for the survival analysis. But
Fig. 1. K–M and fitted survival curves for PPC and baseline Weibull probability density functions. (a) K–M and fitted survival curves for time to PPC by gender and treatment. (b) Curves of baseline Weibull probability density function

\[ f_0(t|k, h) = h_0(t|k, h) \exp[- \int_0^t h_0(u|k, h)du] = k h(t)^{k-1} \exp[-(ht)^k]. \]

Downloaded from https://academic.oup.com/biostatistics/article-abstract/15/3/540/224153 by guest on 10 June 2018
modeling the data using the traditional Cox PH cure model framework does not take advantage of this additional cure information.

In many studies, there are diagnostic procedures available to provide further information about whether a subject is cured. However, the additional information about cured status may not be available for all subjects, and all diagnostic procedures are likely associated with a certain degree of accuracy in terms of sensitivity and specificity. Complete separation of cured and uncured subjects in the censored subset can be difficult to achieve. Hence the PH cure model that incorporates the additional information also needs to take into account the sensitivity and specificity of the diagnostic procedure that produces this additional information.

In this paper, we extend the traditional Cox PH cure models to incorporate the additional diagnostic information about cured status. The proposed extension can be applied when the latency part is modeled by the Cox PH model. Extensive simulations demonstrated that the proposed extension provides more efficient and less biased estimations, and that the higher efficiency and smaller bias was associated with higher sensitivity and specificity of the diagnostic procedures. At the end, the extended Cox PH cure model is applied to the motivating example of the pediatric bone fracture study. We showed that the efficiency gain changed some effects from non-significant to significant after we incorporated the additional cure information into the model.

The paper is organized as follows. In Section 2, we proposed an extension to the traditional Cox PH cure models to incorporate the additional cure information. We also provided parameter estimation procedures of the extended Cox PH cure model. In Section 3, evaluations of the extended cure model were conducted through extensive simulation studies. In Section 4, we applied this extended cure model to the motivating example of the pediatric bone fracture study. Discussion is presented in Section 5.

2. Cox PH cure models with sensitivity and specificity

2.1 Model specification

Consider the data in the form \((t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i), i = 1, 2, \ldots, n\). Here, \(t_i\) denotes the observed survival time for the \(i\)th patient, \(\delta_i\) is the censoring indicator with 1 if \(t_i\) is uncensored (i.e., observed), and 0 if censored. \(\mathbf{x}_i\) and \(\mathbf{z}_i\) are two covariate vectors. Let \(\mathbf{\beta}\) and \(\mathbf{\gamma}\) be the parameter vectors related to \(\mathbf{x}_i\) and \(\mathbf{z}_i\), respectively. If we model this dataset by using the Cox PH cure model specified in (1.1), \(\pi(\mathbf{z}_i) = \exp(\mathbf{\gamma}'\mathbf{z}_i)/[1 + \exp(\mathbf{\gamma}'\mathbf{z}_i)]\), and \(h_u(t_i|\mathbf{x}_i) = h_0(t_i) \exp(\mathbf{\beta}'\mathbf{x}_i)\) and \(S_u(t_i|\mathbf{x}_i) = [S_0(t_i)]^{\exp(\mathbf{\beta}'\mathbf{x}_i)}\) are the hazard function and survival function of uncured subject \(i\), where \(h_0(t)\) and \(S_0(t) = \exp[-\int_0^t h_0(u)\,du]\) are unspecified baseline hazard and survival functions, respectively. Let \(\mathbf{\theta}_0 = (\mathbf{\beta}', \mathbf{\gamma}')\) and \(\mathbf{O}_0 = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i), i = 1, 2, \ldots, n\}\). The observed likelihood for the Cox PH cure model can be written as:

\[
L_{\text{obs}}(\mathbf{\theta}_0; \mathbf{O}_0) = \prod_{i=1}^{n} \left[ \pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i) \right]^{\delta_i} \left[ \pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i) + [1 - \pi(\mathbf{z}_i)] \right]^{1-\delta_i}.
\]

Suppose in addition to \((t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i)\), for censored patients, we also observe the result \(d_i\) from a diagnostic procedure, where \(d_i = 1\) if the \(i\)th patient is diagnosed as cured, 0 if diagnosed as uncured. A diagnostic procedure usually is associated with certain sensitivity and specificity. Sensitivity measures the proportion of actual positives who are correctly identified (e.g., the percentage of sick people who are correctly identified as sick). Specificity measures the proportion of actual negatives who are correctly identified (e.g., the percentage of healthy people who are correctly identified as healthy). Assume that the results of diagnostic procedure are independent of time, i.e. \(d_i\) is independent of \(t_i\), and the diagnostic procedure has a sensitivity of \(p_0\) and a specificity of \(1 - p_1\). For a validated diagnostic procedure, we will have \(p_0 \geq p_1\). We might model \(p_0\) and \(p_1\), but for simplicity, we assume that these conditional probabilities do not depend on any
covariates. Denote \( O_1 = \{(t_i, \delta_i, x_i, z_i, d_i), i = 1, 2, \ldots, n\} \) and \( \theta' = (\theta_0', p_0, p_1) \). The observed likelihood then becomes

\[
L_{\text{obs}}(\theta; O_1) = \prod_{i=1}^{n} \left[ \pi(z_i) f_u(t_i|x_i) \right]^{\delta_i} \{p_1 d_i (1 - p_1)^{1-d_i} \pi(z_i) S_0(t_i|x_i) \}
+ p_0^{d_i} (1 - p_0)^{1-d_i} [1 - \pi(z_i)]^{1-\delta_i}
\]

(2.2)

because for uncensored patients (\( \delta_i = 1 \)), the contribution to the likelihood is the same as that in (2.1); while for censored patients (\( \delta_i = 0 \)), with the independence assumption of \( d_i \) and \( t_i \), the contribution is

\[ p_1^{d_i} (1 - p_1)^{1-d_i} \pi(z_i) S_0(t_i|x_i) \]

if they are uncured, and the contribution is

\[ p_0^{d_i} (1 - p_0)^{1-d_i} [1 - \pi(z_i)] \]

if they are cured.

Because the diagnostic procedure results are not available for all the censored subjects, let \( \eta_i = 1 \) denote the diagnostic result available for subject \( i \), and \( \eta_i = 0 \) denote the diagnostic result unavailable for subject \( i \). Let \( O = \{(t_i, \delta_i, x_i, z_i, \eta_i, d_i), i = 1, 2, \ldots, n\} \), the observed likelihood can be written as follows:

\[
L_{\text{obs}}(\theta; O) = \prod_{i=1}^{n} \left[ \pi(z_i) f_u(t_i|x_i) \right]^{\delta_i}
\times \{p_1^{d_i} (1 - p_1)^{1-d_i} \pi(z_i) S_0(t_i|x_i) + p_0^{d_i} (1 - p_0)^{1-d_i} [1 - \pi(z_i)] \}^{(1-\delta_i)\eta_i}
\times \{\pi(z_i) S_0(t_i|x_i) + [1 - \pi(z_i)] \}^{(1-\delta_i)(1-\eta_i)}.
\]

(2.3)

Notice that when \( p_1 = 0 \) or 1, it implies that the uncured patient \( i \) is either correctly identified as uncured if \( p_1 = 0 \), hence \( d_i = 0 \), or incorrectly identified as cured if \( p_1 = 1 \), hence \( d_i = 1 \). Thus the term \( p_1^{d_i} (1 - p_1)^{1-d_i} \) reduces to 1 in (2.3). Similarly, when \( p_0 = 0 \) or 1, the term \( p_0^{d_i} (1 - p_0)^{1-d_i} \) reduces to 1 in (2.3). Notice that when \( p_0 = p_1 \), (2.3) reduces to (2.1) except for a constant multiplier, meaning that if both sensitivity and (1 - specificity) are the same, the likelihood functions with and without the diagnostic information are the same. In practice, we want both sensitivity and specificity to be high, and \( p_0 \neq p_1 \).

Logistic regression is used to model the “incidence” part \( \pi(z) \) of the mixture model. The Cox PH method is used to model the “latency” part \( S_0(t|x) \) of the mixture model. We use the expectation-maximization (EM) algorithm to estimate the model parameters in (2.3) as Peng did (Peng, 2003a). The following section provides the details of the EM procedure.

### 2.2 Estimation with EM algorithm

Let \( c_i \) be the indicator of the uncured status for subject \( i \), i.e. 1 if the subject is uncured (susceptible), and 0 otherwise, where \( c_i \) is a latent variable but it is partially observed because \( \delta_i = 1 \) implies \( c_i = 1 \). We have:

\[
d_i | (c_i = 0, \delta_i = 0) \sim \text{Bernoulli}(p_0),
\]

\[
d_i | (c_i = 1, \delta_i = 0) \sim \text{Bernoulli}(p_1),
\]
and the complete log-likelihood can be written as:

\[
\ell_c(\theta; O, c) = \log L_c(\theta; O, c) = \sum_{i=1}^{n} \left\{ c_i \log \pi(z_i) + (1 - c_i)(1 - \delta_i) \log[1 - \pi(z_i)] \right\} \\
+ \sum_{i=1}^{n} \left[ c_i \delta_i \log f_u(t_i | x_i) + c_i (1 - \delta_i) \log S_u(t_i | x_i) \right] \\
+ \sum_{i=1}^{n} [d_i \log p_1 + (1 - d_i) \log(1 - p_1)] c_i (1 - \delta_i) \eta_i \\
+ \sum_{i=1}^{n} [d_i \log p_0 + (1 - d_i) \log(1 - p_0)] (1 - c_i) (1 - \delta_i) \eta_i, \\
\]

where \( c = \{ c_i, i = 1, \ldots, n \} \). Notice that \((1 - c_i)(1 - \delta_i) = 1 - c_i \) and \( c_i \delta_i = \delta_i \). (2.4) can be further simplified to:

\[
\ell_c(\theta; O, c) = \sum_{i=1}^{n} \left\{ c_i \log \pi(z_i) + (1 - c_i) \log[1 - \pi(z_i)] \right\} \\
+ \sum_{i=1}^{n} \left[ \delta_i \log h_u(t_i | x_i) + c_i \log S_u(t_i | x_i) \right] \\
+ \sum_{i=1}^{n} [d_i \log p_1 + (1 - d_i) \log(1 - p_1)] c_i (1 - \delta_i) \eta_i \\
+ \sum_{i=1}^{n} [d_i \log p_0 + (1 - d_i) \log(1 - p_0)] (1 - c_i) (1 - \delta_i) \eta_i, \\
\]

where \( h_u(\cdot) = f_u(\cdot)/S_u(\cdot) \).

It can be seen from (2.5) that the complete log-likelihood function can be separated into three parts: the first part contains only the “incidence” parameter vector \( y \) related to the covariate vector \( z \), denoted by \( \ell_{c1}(y; O, c) \), the second part contains only the “latency” parameter vector \( \beta \) related to the covariate vector \( x \), denoted by \( \ell_{c2}(\beta; O, c) \), and the third part contains only the sensitivity parameter \( p_0 \) and specificity parameter \( 1 - p_1 \), denoted by \( \ell_{c3}(p_0, p_1; O, c) \). The three parts can be maximized separately given \( c \). The EM algorithm can be carried out in the following steps.

**Initial value**: The EM algorithm starts with an initial value \( \theta^{(0)} \).

**E-step**: The E-step in the \((r + 1)\)th iteration is used to calculate the expectation of the complete log-likelihood function \( \ell_c(\theta; O, c) \), conditional on the observed data and \( \theta^{(r)} \) the estimate of \( \theta \) at the \( r \)th iteration. This is equivalent to calculating the following conditional expectation:

\[
g^{(r)}_i = E(c_i | \theta^{(r)}, O) = P(c_i = 1 | \theta^{(r)}, O),
\]
which is the $r$th estimator of the uncured probability of the $i$th patient. Because

$$P(c_i = 1|d_i = 1, \delta_i = 0, \theta^{(r)}, \mathbf{O}) = \eta_i \frac{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i) p_1^{(r)}}{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i) p_1^{(r)} + [1 - \pi^{(r)}(z_i)] p_0^{(r)}}$$

$$+ (1 - \eta_i) \frac{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i) [1 - \pi^{(r)}(z_i)]}{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i) + [1 - \pi^{(r)}(z_i)]},$$

and

$$P(c_i = 1|d_i = 0, \delta_i = 0, \theta^{(r)}, \mathbf{O}) = \eta_i \frac{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i)(1 - p_1^{(r)})}{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i)(1 - p_1^{(r)}) + [1 - \pi^{(r)}(z_i)](1 - p_0^{(r)})}$$

$$+ (1 - \eta_i)(1 - d_i) \frac{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i)(1 - p_1^{(r)})}{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i) + [1 - \pi^{(r)}(z_i)]}.$$  

$g_i^{(r)}$ can be expressed as follows:

$$g_i^{(r)} = P(c_i = 1|\theta^{(r)}, \mathbf{O})$$

$$= \delta_i + (1 - \delta_i)d_i \eta_i \frac{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i) p_1^{(r)}}{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i) p_1^{(r)} + [1 - \pi^{(r)}(z_i)] p_0^{(r)}}$$

$$+ (1 - \delta_i)(1 - d_i) \eta_i \frac{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i)(1 - p_1^{(r)})}{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i)(1 - p_1^{(r)}) + [1 - \pi^{(r)}(z_i)](1 - p_0^{(r)})}$$

$$+ (1 - \delta_i)(1 - \eta_i) \frac{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i)}{\pi^{(r)}(z_i) S_u^{(r)}(t_i|x_i) + [1 - \pi^{(r)}(z_i)]}. \quad (2.6)$$

**M-step:** The M-step in the $(r + 1)$th iteration is used to maximize the expected complete log-likelihood function with respect to $\theta$ to obtain $\theta^{(r+1)}$, where the expected complete log-likelihood function is the sum of the following 3 functions:

$$\tilde{\ell}_{c1}(\gamma|g^{(r)}, \mathbf{O}) = \sum_{i=1}^{n} \{g_i^{(r)} \log[\pi(z_i)] + (1 - g_i^{(r)}) \log[1 - \pi(z_i)]\}, \quad (2.7)$$

$$\tilde{\ell}_{c2}(\beta|g^{(r)}, \mathbf{O}) = \sum_{i=1}^{n} \{\delta_i \log[h_u(t_i|x_i)] + g_i^{(r)} \log[S_u(t_i|x_i)]\}, \quad (2.8)$$

$$\tilde{\ell}_{c3}(p_0, p_1|g^{(r)}, \mathbf{O}) = \sum_{i=1}^{n} \{d_i \log p_1 + (1 - d_i) \log(1 - p_1)]g_i^{(r)}(1 - \delta_i)\eta_i$$

$$+ \sum_{i=1}^{n} \{d_i \log p_0 + (1 - d_i) \log(1 - p_0)](1 - g_i^{(r)})\eta_i, \quad (2.9)$$

where $g^{(r)} = \{g_i^{(r)}, i = 1, \ldots, n\}$.

Equation (2.7) is the log-likelihood function of the logistic regression model for values arising from a binomial distribution with the response probability $\pi(z_i) = \exp(\gamma z_i)/[1 + \exp(\gamma z_i)]$. It can be maximized by the usual optimization methods such as the Newton–Raphson method. This maximization procedure can be carried out in most standard logistic regression programs to obtain the estimate of $\gamma$. 
The maximization of \( \ell_{c2}(\beta | g^{(r)}, O) \) in (2.8) can be handled by using the method of Peng (2003a), i.e. we first obtain the estimate of \( \beta \) by a program for the Cox PH model that accepts covariates with fixed coefficients, and after getting \( \beta^{(r+1)} \), the estimate of the baseline survival function \( S_0(t) \), denoted by \( S_0^{(r+1)}(t) \), can be obtained semi-parametrically (Peng, 2003b). Peng (2003b) proposed several methods of tail completion for the estimation of baseline survival function. We found that the zero tail completion already provided a satisfactory result in our simulation settings. It may be worthwhile to use more advanced tail completion methods when there is a need in more complicated situation.

Lastly, we maximize \( \ell_{c3}(p_0, p_1|g^{(r)}, O) \) in (2.9) to obtain \( p_0^{(r+1)} \) and \( p_1^{(r+1)} \) explicitly by using the following updating formula:

\[
P_0^{(r+1)} = \frac{\sum_{i=1}^{n} d_i \eta_i (1 - g_i^{(r)}) (1 - \delta_i)}{\sum_{i=1}^{n} \eta_i (1 - g_i^{(r)}) (1 - \delta_i)} = \frac{\sum_{i: \delta_i = 0 \& \eta_i = 1} d_i (1 - g_i^{(r)})}{\sum_{i: \delta_i = 0 \& \eta_i = 1} (1 - g_i^{(r)})}, \tag{2.10}
\]

\[
P_1^{(r+1)} = \frac{\sum_{i=1}^{n} d_i \eta_i g_i^{(r)} (1 - \delta_i)}{\sum_{i=1}^{n} \eta_i g_i^{(r)} (1 - \delta_i)} = \frac{\sum_{i: \delta_i = 0 \& \eta_i = 1} d_i g_i^{(r)}}{\sum_{i: \delta_i = 0 \& \eta_i = 1} g_i^{(r)}}, \tag{2.11}
\]

Notice that if all available \( d_i \)s for censored subjects are 1, the estimates of \( p_0^{(r+1)} \) and \( p_1^{(r+1)} \) are 1; similarly, if all available \( d_i \)s for censored subjects are 0, the estimates of \( p_0^{(r+1)} \) and \( p_1^{(r+1)} \) are 0. If the sensitivity and specificity are known externally from the diagnostic procedure, the estimations of \( p_0 \) and \( p_1 \) are not needed. But caution must be taken when using the external sensitivity and specificity values because it might not apply to the observed data well.

Iteration: The algorithm is iterated until \( \| \theta^{(r+1)} - \theta^{(r)} \| \) is sufficiently small.

To estimate the variance of the EM estimators, the method of Louis (1982) may be used. However, as suggested in Peng (2003b), “The accuracy of the approximation is not clear and it cannot be obtained from the iterations of the EM algorithm”. Therefore, we follow the suggestion of Peng (2003b) to estimate the variance of the estimated parameters by using the bootstrap method.

### 3. Evaluations of the extended Cox PH cure model

Extensive simulations were performed to evaluate the performance of the proposed model. Specifically, we compared the following 3 methods: (i) the extended Cox PH cure model incorporating diagnostic information, with unknown sensitivity (\( p_0 \)) and specificity (\( 1 - p_1 \)); (ii) the extended Cox PH cure model incorporating diagnostic information, with \( p_0 \) and \( 1 - p_1 \) known a priori; and (iii) traditional Cox PH cure model without diagnostic information.

To mimic the pediatric bone data, we first generated \( c_i \) according to the incidence model specified by \( \logit(\pi_i) = \gamma_0 + \gamma_1 I_{(TRT=1)} + \gamma_2 I_{(TRT=2)} + \gamma_3 I_{(SEX=Male)} \). Here \( \pi_i \) is the uncured probability of subject \( i \). Covariates TRT and SEX are evenly distributed three-level and two-level covariates. The true parameter values were \( \gamma_0 = 0.25, \gamma_1 = -0.1, \gamma_2 = 0.5, \) and \( \gamma_3 = -0.1 \). Survival data were simulated from the latency part, specified by the PH model: \( h_u(t_i) = h_0(t_i) \exp\{\beta_1 I_{(TRT=1)} + \beta_2 I_{(TRT=2)} + \beta_3 I_{(SEX=Male)}\} \), where the true parameter values were \( \beta_1 = 0.2, \beta_2 = -0.3, \) and \( \beta_3 = 0.1 \). In order to compare the performance of these 3 methods under varying censoring scenarios, we parameterized the baseline hazard function by using the Weibull hazard distribution: \( h_0(t|k, h) = kh(ht)^{k-1} \), and four different sets of \( (h, k) \) were considered:

\[
(1) \ h = 1, k = 2; \quad (2) \ h = 2, k = 2; \quad (3) \ h = \frac{3}{2}, k = 3; \quad \text{and} \quad (4) \ h = \frac{1}{3}, k = 4. \tag{3.1}
\]

We set the maximum follow-up time at 6 and generated censoring times from the uniform(0, 6) distribution. The baseline Weibull pdf is \( f_0(t|k, h) = h_0(t|k, h) \exp[-\int_0^t h_0(u|k, h)du] = kh(ht)^{k-1} \exp[-(ht)^k] \). The
shapes of these four baseline Weibull pdf under the four parameter sets in (3.1) are shown in Figure 1(b). After $c_i$ and the censoring status were determined, we generated $d_i$ based on the following Bernoulli distributions: $d_i | (c_i = 0, \delta_i = 0) \sim \text{Bernoulli}(p_0)$, and $d_i | (c_i = 1, \delta_i = 0) \sim \text{Bernoulli}(p_1)$, with true sensitivities $p_0$ of 30%, 50%, 70% and 100%, and true specificities $1 - p_1$ of 80%, 90% and 100%. Simulations with 100% of subjects having available diagnostic information were performed for all settings described above. For each simulation scenario, 200 subjects were simulated, and a total of 1,000 simulation runs were performed. Variances of parameter estimates were estimated based on 1,000 bootstrap samples. Based on simulated data, the average censoring rates of the four parameter sets were 50.5% when $h = 1$, $k = 2$; 46.3% when $h = 2$, $k = 2$; 55.0% when $h = \frac{2}{3}$, $k = 3$; 68.3% when $h = \frac{1}{3}$, $k = 4$.

We evaluated the performance of regression parameter estimates in regard to biasedness, mean squared error (MSE) and relative efficiency (RE) to the traditional method, with results summarized in Figures 2, 3, and 4, respectively. In the graphs, “CL” stands for the traditional Cox PH cure model, subscript $u$ under specificity values indicates unknown sensitivity and specificity (Rows 1, 3 and 5), and numbers 1 to 4 correspond to the four parameter sets of the baseline Weibull hazard distribution in (3.1).

Overall, the point estimates of all three methods were close to the true parameter values because the biases were small (see Figure 2), indicating the consistency of parameter estimates. For the methods incorporating diagnostic information, MSEs decreased with both sensitivity $p_0$ and specificity $1 - p_1$ (see Figure 3), and the MSEs were consistently smaller than those of the traditional method. Moreover, the decrease in MSEs when sensitivity and specificity are unknown was smaller than those when sensitivity and specificity are known. In the extended models, the RE to the traditional method increased with both sensitivity $p_0$ and specificity $1 - p_1$ (see Figure 4). The increase in RE when sensitivity and specificity are unknown was smaller than that when sensitivity and specificity are known. As expected, the improvements in RE for incidence parameter estimates are bigger than those for latency parameter estimates. The greatest improvement in RE was achieved when the censoring rate was high [e.g., censoring rate of 68.3% for $h = \frac{1}{3}$ and $k = 4$ (Parameter set 4), followed by 55.0% for $h = \frac{2}{3}$ and $k = 3$ (Parameter set 3), 50.5% for $h = 1$ and $k = 2$ (Parameter set 1), and 46.3% for $h = 2$, $k = 2$ (Parameter set 2)]. It suggested that incorporating additional diagnostic information can substantially alleviate the ambiguity in subjects’ cured status and improve efficiency of parameter estimates.

To investigate the impact on the estimation when not all subjects have available diagnostic information, we performed simulations where only 50% of subjects had diagnostic information, and compared these simulations to those where 100% of subjects had diagnostic information. Data were generated in the same fashion as described previously. However, we only simulated data with $h = \frac{1}{3}$ and $k = 4$ for the baseline Weibull hazard function in the latency part. Bottom rows of Figures 2, 3, and 4 summarize the biasedness, MSEs, and REs for the estimated regression parameters under the following scenarios: (a) 100% of subjects having diagnostic information, and with known $p_0$ and $1 - p_1$; (b) 50% of subjects having diagnostic information, and with known $p_0$ and $1 - p_1$; (c) 100% of subjects having diagnostic information, and with unknown $p_0$ and $1 - p_1$, and (d) 50% of subjects having diagnostic information, and with unknown $p_0$ and $1 - p_1$.

By comparing scenarios (a) vs. (b) or scenarios (c) vs. (d), it is obvious that MSEs were greater and REs were smaller when fewer subjects had available diagnostic information. However, with either known or unknown $p_0$ and $1 - p_1$, MSEs were consistently smaller and REs were consistently higher if additional diagnostic information was incorporated, even it was only available for 50% of subjects.

4. Application to Pediatric Bone Data

This is a clinical study that retrospectively reviewed 157 (75 girls and 82 boys) children’s charts to identify the incidence of PPC following physeal fractures of the distal end of the tibia (Leary and others, 2009).
Among the 157 children, there were 125 (79.6%) children treated with Cast. Sixteen out of the 157 children were identified as having PPC. Other children were considered to be cured if the symmetric Harris growth arrest line was observed or closure of the growth plate was seen radiographically. Ninety-six children were considered as cured. The diagnostic cured statuses of forty-five children could not be determined, and hence their diagnostic cured statuses were considered unavailable.
Because there appeared a clear cure indication in this data set as shown from the K–M curves of the time to PPC by gender and treatment in Figure 1(a), it is appropriate to use a Cox PH cure model for the survival analysis. A subject’s cured status could be indicated by the symmetric Harris growth arrest line or closure of the growth plate. Because the cure indicator in this case is definitive, it may be treated as a diagnostic procedure with known 100% sensitivity and specificity. We also treated sensitivity and
Standard errors of parameter estimates were estimated based on 1,000 bootstrap samples.

Figure 1(a). The graph shows that the model fitted reasonably well. The factor of treatment methods (Cast and non-Cast) and gender effects were included in the survival portion and the cure portion of the model.

<table>
<thead>
<tr>
<th>Spec=1.0 u</th>
<th>Spec=0.9 u</th>
<th>Spec=0.8 u</th>
<th>Spec=1.0 u</th>
<th>Spec=0.9 u</th>
<th>Spec=0.8 u</th>
<th>Spec=1.0 u</th>
<th>Spec=0.9 u</th>
<th>Spec=0.8 u</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
</tr>
<tr>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td>Spec=0.8 u</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td>Spec=0.9 u</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td>Spec=1.0 u</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4. Relative efficiency.

Sensitivity

known & all available: ———
known but only 50% available: ———
unknown & all available: ———
unknown but only 50% available: ———

specificity as unknown in our data analysis and showed fitted survival curves by gender and treatment in Figure 1(a). The graph shows that the model fitted reasonably well. The factor of treatment methods (Cast and non-Cast) and gender effects were included in the survival portion and the cure portion of the model. Standard errors of parameter estimates were estimated based on 1,000 bootstrap samples.
Table 1. Application of Cox PH cure model with and without diagnostic information to pediatric bone data

<table>
<thead>
<tr>
<th>Survival portion</th>
<th>Classical PH cure model</th>
<th>Extended PH cure model with known $p_0$ and $p_1$</th>
<th>Extended PH cure model with unknown $p_0$ and $p_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
<td>log(HR)</td>
<td>SE</td>
<td>$P$-value</td>
</tr>
<tr>
<td>Male</td>
<td>−0.479</td>
<td>3.245</td>
<td>0.8827</td>
</tr>
<tr>
<td>Cast</td>
<td>−0.447</td>
<td>2.602</td>
<td>0.8636</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Logistic portion</th>
<th>Effect</th>
<th>log(OR)</th>
<th>SE</th>
<th>$P$-value</th>
<th>log(OR)</th>
<th>SE</th>
<th>$P$-value</th>
<th>log(OR)</th>
<th>SE</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−0.536</td>
<td>1.321</td>
<td>0.6850</td>
<td>−1.189</td>
<td>0.517</td>
<td>0.0214</td>
<td>−0.357</td>
<td>0.372</td>
<td>0.3375</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>−0.610</td>
<td>1.506</td>
<td>0.6865</td>
<td>−0.337</td>
<td>0.475</td>
<td>0.4786</td>
<td>−0.107</td>
<td>0.248</td>
<td>0.6667</td>
<td></td>
</tr>
<tr>
<td>Cast</td>
<td>−0.155</td>
<td>1.438</td>
<td>0.9142</td>
<td>−0.887</td>
<td>0.514</td>
<td>0.0841</td>
<td>−0.131</td>
<td>0.344</td>
<td>0.7036</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 compares the results of applying the classical Cox PH cure model and the extended PH cure models with known and unknown sensitivity and specificity. All models provide the same sign and similar magnitude for the point estimates, while the estimated standard errors from the extended models with the diagnostic information incorporated are much smaller. This observation is consistent with our simulation results. A more efficient estimation leads to a more powerful test and might lead to a different conclusion. As shown in the table, the 2-sided $P$-value for logistic intercept is 0.6850 based on the classical PH cure model; this is changed to a significant $P$-value of 0.0214 for the extended PH cure model with known sensitivity and specificity. This is consistent with the pattern of the data: we have 96 subjects that were cured, so the incidence part of the model can not be ignored. A similar observation is also noted for the Cast factor in the logistic portion: the 2-sided $P$-value is 0.9142 from the classical PH cure model, but becomes marginally significant ($P$-value = 0.0841) from the extended PH cure model. Notice that the choice of the external information of sensitivity and/or specificity may affect the point estimates. Table 1 also shows that the estimated standard errors from the extended PH cure model with unknown sensitivity and specificity are smaller than those from that with known sensitivity and specificity. Investigation of the data reveals that this is due to the fact that all the censored subjects are all declared as either cured or unknown, and we do not have censored subjects declared as uncured. Based on the ML estimates from the model (see updating formulas (2.10) and (2.11)), $p_0$ and $p_1$ were always 1 in each iteration step and hence it was the same as if they were known. The results show that using the ML estimates performs better than using the external sensitivity and specificity values. One interpretation is that the external sensitivity and specificity values might not be applicable to the observed data well when all of the censored patients having cure information were declared as cured or all of them were declared as uncured. Further research in this direction might be needed in the future.

5. Discussion

In a recent issue of a medical journal (Othus and others, 2012), cure models are advocated for survival data analysis when there is evidence of long-term survivors. Traditional cure models assume that the cured or uncured status in the censored set can not be distinguished. In this paper, a novel extension of a Cox parametric PH cure model is proposed. With the inclusion of additional diagnostic information, the cured or uncured status can be partially distinguished to the extent of certain sensitivity and specificity. The extended Cox PH cure model has the potential of improving the efficiency of the estimation.
In this paper, extensive simulations are performed for evaluations of the extended Cox PH cure model. The simulations show that the proposed extension provides more efficient and less biased estimations. All the biases are small, MSEs are smaller when additional diagnostic procedures are incorporated, and relative efficiencies are greater than 1 when additional diagnostic procedures are incorporated. Higher sensitivity and specificity are associated with smaller MSE and larger RE. The larger efficiency gain is noted when the censoring rate is high. When the censoring rate is high, the set of subjects with uncured or cured status undetermined is larger; hence, adding additional diagnostic information provides a larger degree of efficiency gain. Based on all the simulation results, the additional cure information, even with availability of only a portion of subjects, incorporated into the Cox PH cure model improves the efficiency of the estimation and provides a less biased estimation and smaller MSE of the estimation. Furthermore, the application of the extended Cox PH cure model to our motivating example shows that there is a significant efficiency gain that results in some effects changing from non-significant to significant. In conclusion, based on the results provided in this paper, we highly recommend that when there is additional cure information available, even only partially, we should incorporate that information into the model. It is also a good idea that we devise diagnostic procedures of cure and collect the available cure information when we design and conduct studies.

Acknowledgments

The authors are grateful to the two referees and AE for their helpful comments that improved the presentation. Conflict of Interest: None declared.

Funding

The research of Y.L., S.L., and W.J.S. was partially supported by NIH/NCI CCSG Grant 3P30CA072720. The research of C.S.L. was partially supported by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH), through grant #UL1 TR000002.

References


[Received April 22, 2013; revised December 30, 2013; accepted for publication December 31, 2013]