Published in final edited form as:

J Am Stat Assoc. 2007 December 1; 102(480): 1235–1244. doi:10.1198/016214506000001446.

Inference of Tamoxifen's Effects on Prevention of Breast Cancer from a Randomized Controlled Trial

Yu Shen [Professor],

Department of Biostatistics, M. D. Anderson Cancer Center, The University of Texas, Houston, TX 77030, Phone: 713-794-4159, Fax: 713-745-4940

Jing Qin [Mathematical Statistician], and

Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892

Joseph P. Costantino [Professor]

Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15213

Yu Shen: yshen@mdanderson.org; Jing Qin: jingqin@niaid.nih.gov; Joseph P. Costantino: costan@nsabp.pitt.edu

SUMMARY

The largest randomized, double-blind, placebo-controlled chemoprevention trial, the National Surgical Adjuvant Breast and Bowel Project's Breast Cancer Prevention Trial (NSABP-BCPT), evaluated the efficacy of tamoxifen in the prevention of breast cancer among women at high risk of developing the disease. The trial has reported a reduction of breast cancer incidence for the tamoxifen group. However, the effect of tamoxifen on the time to diagnosis of the disease over the six-year follow-up of the trial has not been fully explored in literature. We propose a flexible semiparametric model to assess the effects of tamoxifen on the incidence of breast cancer as well as time to the diagnosis of the disease, separately, in the framework of a cure-rate model. We used an estimating equation approach to estimating the unknown parameters, and assessed the semiparametric model assumption with a test based on the area between two survival curves. On the NSABP-BCPT data, we found that the treatment of tamoxifen has a substantial effect in the reduction of invasive breast cancer events in estrogen receptor (ER)-positive tumors, but has no effect on ER-negative tumors. Among women who were diagnosed to have ER-positive breast cancer during the study follow-up, there was little difference in terms of time to diagnosis between the two arms. However, tamoxifen may advance the time to breast cancer diagnosis for ER-negative breast cancer, while the incidence of ER-negative tumors is similar in the two treatment arms.

Keywords

Cancer prevention; Cure rate model; NSABP trial; Semiparametric model; Tamoxifen

1 Introduction

The National Surgical Adjuvant Breast and Bowel Project's Breast Cancer Prevention Trial is the largest randomized, double-blind, placebo-controlled clinical trial to evaluate the efficacy of tamoxifen (TAM) in the prevention of breast cancer among women at high risk of developing the disease (Fisher et al, 1998). A total of 13,388 asymptomatic participants aged 35 years and older were equally randomized to receive either placebo or tamoxifen daily for 5 years. Participants were monitored for development of the disease during the six years of follow-up.

The study lead to the conclusion that tamoxifen substantially decreased the incidence of invasive breast cancer among the study cohort (Fisher, et. al, 1998 and Fisher, et. al, 1999).

While it is clear that the incidence of breast cancer in the treatment group was reduced by 49% compared to the placebo group, it is unclear how tamoxifen may have affected women who were diagnosed with breast cancer. Specifically, whether tamoxifen treatment shifted the time of diagnosis of breast cancer among the diseased cases. Moreover, clinical trials have shown clear benefits of TAM for patients with ER-positive breast cancer. However, no significant treatment effect has been found for ER-negative breast cancer (Cuzick et al, 2003; EBCTCG, 2005; Kaufmann et al, 2006). It is thus of great interest to explore the treatment effects of TAM within sub-cohorts of patients as determined by the ER status of their tumors, both on the disease incidence and the diagnosis of breast cancer.

Using existing survival analysis methods to analyze cancer prevention data is attractive, as the primary endpoint of interest is the diagnosis of breast cancer, which may be right-censored for most participants due to a low incidence of the disease and relatively short follow-up. The study cohort is a mixture of individuals, some of whom are likely to develop breast cancer in their lifetime while others are not. Cure models thus become more relevant in this setting. In such a framework, we may investigate the treatment effect on the disease incidence and the effect on the change of time to the diagnosis, separately. Various parametric and semiparametric models have been used for the purpose of testing and estimation with right-censored survival data (Farewell 1982; Maller and Zhou 1992; Tsodikov 1998; Chen et al 1999; Broet et al 2001; Ibrahim et al, 2001a,b; Sy and Taylor 2002; Tsodikov 2002; Tsodikov et al 2003). However, their applications in cancer prevention trials have been limited. Models other than the proportional hazards model are required to describe the treatment effects using cure models in the application of cancer prevention studies (Tsodikov, 2002). Our goals are therefore to estimate the distribution of time to the diagnosis of breast cancer, and the incidence of the disease within each intervention group, and to separately assess the treatment effects of TAM for preventing the disease and for changing time to the diagnosis of breast cancer.

In the BCPT trial, let T be the time to the disease diagnosis, U be the time to the last follow-up, $X = \min(T, U)$ and $\delta = I(T \leq U)$ be the indicator for the event of interest, and $Z = 1$ for tamoxifen treatment and 0 for placebo. Assume that the censoring variable U is independent of T given Z . In the k th arm, assume that $\pi_k = pr(D = 1|Z = k) = pr(T < \infty|Z = k)$ is the probability that the disease would occur if the follow-up time were not right-censored, where $D = 1$ indicates the disease status and $k = 0$ or 1. In the following text, this probability will be referred to as *cumulative incidence* or *incidence* for simplicity.

The survival distribution of T in the control arm follows a mixture model

$$\bar{F}(t) = \pi_0 \bar{F}_c(t|D=1) + (1 - \pi_0),$$

where $1 - \pi_0$ is often referred to as the proportion of “non-susceptible” cases in the cohort. The survival function, $\bar{F}(t)$, is an improper distribution of the time to disease diagnosis, and $\bar{F}_c(t|D = 1)$ is the proper survival function for the time to disease diagnosis, conditional upon the ultimate occurrence of the disease. Subsequently, the corresponding probability density functions have the following relationship $f(t) = \pi_0 f_c(t)$, where f_c is the probability density function (pdf) of T conditional on $D = 1$, and f is the pdf of T .

In the TAM treatment arm, the unconditional survival distribution of time to the disease diagnosis then follows

$$\bar{G}(t) = \pi_1 \bar{G}_c(t|D=1) + (1 - \pi_1),$$

where $G^-(\cdot)$ is the survival distribution of T . The pdf and cdf of time to the disease diagnosis conditional on $D = 1$ in the TAM group are defined by g_c and $G_c = 1 - \bar{G}_c$.

The mixture model allows us to separately estimate the distribution of the time to the diagnosis among the diseased cases, as well as the disease incidence within each treatment arm.

Furthermore, we want to assess the treatment effect on both the incidence and the time to the diagnosis by relating the data from the two treatment arms through a semiparametric structure. The nonparametric discrete density curves for the time to disease diagnosis show similar shapes but with different scales for the two treatment arms (not shown here due to the limited space). Since the curves do not imply an obvious parametric model for the data, a flexible model without specifying any parametric forms for the densities would be desirable.

Among all diseased cases in the study cohort, we quantify the potential change in the time to the disease diagnosis between the treatment and placebo arms by

$$g_c(t) = \exp\{\alpha^* + \beta h(t)\} f_c(t), \quad g(t) = \frac{\pi_0}{\pi_1} \exp\{\alpha^* + \beta h(t)\} f(t), \quad (1.1)$$

where h is a specified function to transform the observed time t , and g_c and f_c are unspecified positive density functions. The above semiparametric model is flexible, in which only the ratio of the two density functions is assumed to have a parametric form. When $h(t) = t$, the model is equivalent to a semiparametric model in which the log ratio of two density functions is linear, or referred to as an exponential tilt model (Anderson, 1979). When $h(t) = \log(t)$ and $\beta = 1$, it is a length-biased model which has often been used to analyze the onset of disease in observational studies (Cox, 1969, Zelen and Feibleib, 1969, Simon, 1980, Vardi, 1985 and Wang, 1996). When $0 < \beta < 1$, model (1.1) describes a weaker change between the two density functions than the length-biased sampling (Qin et al, 2002), so that the length biased pdf is proportional to the power β of its length. We will refer to the model with $0 < \beta \leq 1$ as the generalized length-biased model. Kay & Little (1987) discussed various choices of the density ratio model for some members of the exponential family.

The natural connection between model (1.1) and the logistic regression model further makes it a popular choice in applications. Conditional on $D = 1$, observe that

$$\text{pr}(Z=1|t) = \frac{\exp(\alpha_1 + \beta h(t))}{1 + \exp(\alpha_1 + \beta h(t))}, \quad \alpha_1 = \alpha^* - \log\{\text{pr}(Z=0)/\text{pr}(Z=1)\}.$$

Some recent applications have shown that model (1.1) with a proper choice of h provides a good fit to the data (Qin & Zhang, 1997; Qin et al., 2002; White & Thompson, 2003 and Qin and Leung, 2005), whereas conventional parametric distributions may not fit the observed data well. Although this class of semiparametric models is robust and popular recently, there is limited work for right-censored data with a cure rate model.

With the observed data from both arms, let $x_1, \dots, x_{n_0}$ be the observed time to the disease diagnosis or censoring in the placebo arm, $x_{n_0+1}, \dots, x_n$ ($n = n_0 + n_1$) be the time to the disease diagnosis or censoring in the TAM group, and δ_i be the corresponding event indicator and $i = 1, \dots, n$. The full likelihood function is proportional to

$$L = \prod_{i=1}^{n_0} \{\pi_0 f_c(x_i)\}^{\delta_i} \{\pi_0 \bar{F}_c(x_i) + 1 - \pi_0\}^{(1-\delta_i)} \prod_{j=n_0+1}^n \{\pi_1 g_c(x_j)\}^{\delta_j} \{\pi_1 \bar{G}_c(x_j) + 1 - \pi_1\}^{(1-\delta_j)}.$$

In general, estimating the unknown parameters from the full likelihood function requires the EM algorithm, which can be computationally intensive when the incidence of the disease is low. Under the semi-parametric model (1.1) in the current setup, it is not clear how to obtain the maximum likelihood estimators based on the full likelihood.

In Section 2, we will describe a conditional estimating equation approach that generalizes the profile likelihood method developed by Qin and Zhang (1997) to the setting with censored survival data. The corresponding survival distributions and the incidence of the disease for each arm can be estimated consistently. In Section 3, the semiparametric model assumption is assessed by using a test based on the area between two survival curves. We also describe two test statistics for assessing the treatment effect on the time to the diagnosis, and the effect on the cumulative incidence of the disease. In Section 4, we present simulation results for the proposed methods. In Section 5 we apply the methods to analyze the NSABP-BCPT trial data, and in Section 6 we provide some concluding remarks.

2 Models

2.1 Conditional on the Diagnosed Cases

An alternative method to the full likelihood is based on the estimating equations conditional on the observed cases only ($\delta = 1$). Let f_1 define the pdf of T conditional on $\delta = 1$ in the placebo arm, and g_1 be pdf of T conditional on $\delta = 1$ in the TAM arm. The true disease status $D = 1$ for all trial participants may or may not be observed due to right-censoring of the follow-up. Therefore, $\{\delta = 1\} \subset \{D = 1\}$. The conditional density of T among the diagnosed cases in the placebo arm then follows

$$f_1(t) = \text{pr}(T=t|\delta=1) = \frac{\text{pr}(T=t, U>t)}{\text{pr}(T<U)}.$$

Under the independent censoring assumption, we can express the above formula as

$$f_1(t) = \frac{\pi_0 f_c(t) \bar{H}_0(t)}{\int_0^\infty \pi_0 f_c(v) \bar{H}_0(v) dv} = \frac{f_c(t) \bar{H}_0(t)}{\mu(f_c, H_0)},$$

where $H_0(t)$ is the “survival” function of the censoring variable U in the placebo arm, and $\mu(f_c, H_0) = \int_0^\infty f_c(v) \bar{H}_0(v) dv$ is a constant with unknown quantities for f_c and H_0 . Similarly, the pdf of T given $\delta = 1$ in TAM is defined by

$$g_1(t) = \frac{\pi_1 g_c(t) \bar{H}_1(t)}{\int_0^\infty \pi_1 g_c(v) \bar{H}_1(v) dv} = \frac{g_c(t) \bar{H}_1(t)}{\mu(g_c, H_1)},$$

where $H_1(t)$ is the “survival” function of the censoring variable U in the TAM arm.

Let $y_1, \dots, y_{m_0}, y_{m_0+1}, \dots, y_m$ denote the observed failure times (i.e., time from randomization to disease diagnosis) among $x_1, \dots, x_{n_0}, x_{n_0+1}, \dots, x_n$, where $m = m_0 + m_1$, and m_0 and m_1 are

the number of diagnosed breast cancer cases in the placebo and TAM arms, respectively. Under the model assumption (1.1) between f_c and g_c , we can relate f_1 and g_1 by the following equations,

$$\frac{g_1(y)}{f_1(y)} = \frac{g_c(y)\bar{H}_1(y)\mu(f_c, H_0)}{f_c(y)\bar{H}_0(y)\mu(g_c, H_1)} := \exp\{\alpha + \beta h(y) + \psi(y)\}, \quad (2.2)$$

where $\alpha = \alpha^* + \log(\mu(f_c, H_0)) - \log(\mu(g_c, H_1))$, and $\psi(y) = \log H^{-1}_1(y) - \log H^{-1}_0(y)$. Given the observed disease cases, we have the corresponding conditional log-likelihood, by generalizing the method of Qin and Zhang (1997) to censored survival data,

$$\begin{aligned} \ell_c(\alpha, \beta, F_c) &= \sum_{i=1}^{m_0} \log f_1(y_i) + \sum_{j=m_0+1}^m \{\log f_1(y_j) + \alpha + \beta h(y_j) + \psi(y_j)\} \\ &:= \sum_{i=1}^m \log p_i + \sum_{j=m_0+1}^m \{\alpha + \beta h(y_j) + \psi(y_j)\} \end{aligned}$$

subject to the constraint that

$$\sum_{i=1}^m p_i = 1, \quad p_i \geq 0, \quad \text{and} \quad \sum_{i=1}^m p_i \{\exp(\alpha + \beta h(y_i) + \psi(y_i)) - 1\} = 0,$$

where $p_i = f_1(x_i)$. Note that the two constraints are to guarantee $\int f_1(x)dx = 1$ and $\int g_1(y)dy = 1$. If the censoring distributions of the two arms are the same, $\psi(y) = 0$. Maximizing the logistic likelihood is equivalent to solving the score (estimating) function in ordinary logistic regression, which would produce consistent estimators of α and β .

If $\psi(\neq 0)$ is known, the score equations for (α, β) follow as

$$\begin{aligned} \frac{\partial \ell_c}{\partial \alpha} &= m_1 - \sum_{i=1}^m \frac{\rho \exp(\alpha + \beta h(y_i) + \psi(y_i))}{1 + \rho \exp(\alpha + \beta h(y_i) + \psi(y_i))}, \\ \frac{\partial \ell_c}{\partial \beta} &= \sum_{i=m_0+1}^m h(y_i) - \sum_{i=1}^m \frac{\rho h(y_i) \exp(\alpha + \beta h(y_i) + \psi(y_i))}{1 + \rho \exp(\alpha + \beta h(y_i) + \psi(y_i))}, \end{aligned}$$

where $\rho = m_1/m_0$. In vector form, they can be equivalently expressed as

$$\begin{pmatrix} \frac{\partial \ell_c}{\partial \alpha} \\ \frac{\partial \ell_c}{\partial \beta} \end{pmatrix} = \sum_{i=m_0+1}^m \eta(y_i) - \rho \sum_{i=1}^{m_0} \eta(y_i) \exp\{\alpha + \beta h(y_i) + \psi(y_i)\}, \quad (2.3)$$

where

$$\eta(y) = \begin{pmatrix} 1 \\ h(y) \end{pmatrix} \frac{1}{1 + \rho \exp\{\alpha + \beta h(y) + \psi(y)\}}.$$

By taking the expectation to the right side of equation (2.3), we have

$$(m_1)E_{g_1}\eta(Y) - m_0\rho E_{f_1}[\eta(Y)\exp\{\alpha+\beta h(Y)+\psi(Y)\}]=0,$$

under model (2.2). Therefore, the estimating equation from the score function is unbiased. Assume that f_c and g_c have finite mean, and finite and nonzero variance. In addition, we assume that $0 < E_{f_c}\{h(y)\}^2 < \infty$ and $0 < E_{g_c}\{h(y)\}^2 < \infty$ which would warrant the existence of $E_{f_c}\{m(y)\}^2$ and $E_{g_c}\{m(y)\}$, where $m(y) = h(y) \exp(\alpha + \beta h(y) + \psi(y)) / (1 + \rho \exp(\alpha + \beta h(y) + \psi(y)))$. Under the above mild regularity conditions, the derivative of (2.3) with respect to (α, β) is a positive definite matrix, so the solution is unique. As a result, solving estimating equation (2.3) produces consistent estimators of (α, β) .

In general ψ is unknown. Applying the method similar to Cheng, Wei and Ying (1995), we replace ψ in the estimating equation by its consistent estimator, $\hat{\psi} = \log(H^{\wedge}_1) - \log(H^{\wedge}_0)$ where $H^{\wedge}_1$ and $H^{\wedge}_0$ are the Kaplan-Meier estimators for the survival functions of the censoring variables in the TAM and placebo arms. By using the arguments of asymptotically unbiased estimating equations for (α, β) , we readily obtain the consistent estimators of $(\hat{\alpha}, \hat{\beta})$.

With the estimated $(\hat{\alpha}, \hat{\beta})$, we can subsequently obtain the estimation of $F_1(t) = \int_0^t f_1(v)dv$, denoted by $F^{\wedge}_1(t) = \sum_{i: y_i < t} \hat{p}_i$, from

$$\hat{p}_i = m_0^{-1} \{1 + \rho \exp[\hat{\alpha} + \hat{\beta}h(y_i) + \hat{\psi}(y_i)]\}^{-1},$$

and $G^{\wedge}_1(t) = \sum_{i: y_i < t} \hat{q}_i$ from $q_i = g_1(y_i)$, where $G_1(t) = \int_0^t g_1(v)dv$,

$$\hat{q}_i = \exp(\hat{\alpha} + \hat{\beta}h(y_i) + \hat{\psi}(y_i))\hat{p}_i.$$

2.2 Estimation of Other Related Quantities

To estimate the distribution of the time to the disease diagnosis among all expected cases ($D = 1$) for each treatment arm, f_c and g_c , we need to derive the functional relationship between f_c and f_1 given in Section 2.1,

$$dF_1(t) = \frac{dF_c(t)\bar{H}_0(t)}{\mu(f_c, H_0)}, \quad (2.4)$$

where both f_1 and f_c are proper pdfs. Thus, we have

$$1 = \int_0^\infty dF_c(t) = \mu(f_c, H_0) \int_0^\infty \bar{H}_0^{-1}(t) dF_1(t).$$

As a consequence, the constant μ can be consistently estimated by

$$\hat{\mu}(f_c, H_0) = 1 / \int \hat{\bar{H}}_0^{-1}(t) d\hat{F}_1(t) = \left\{ \sum_{i=1}^m \frac{\hat{p}_i}{\hat{\bar{H}}_0(y_i)} \right\}^{-1},$$

where $\hat{H}_0(y_i) = 1 - \hat{H}_0(y_i)$ is obtained from the Kaplan-Meier estimator. From (2.4), $F_c(t)$ can then be estimated by

$$\hat{F}_c(t) = \hat{\mu}(f_c, H_0) \sum_{i=1}^m \hat{H}_0^{-1}(y_i) \hat{p}_i I(y_i \leq t).$$

Under the model (1.1), $G_c(t)$ can be estimated similarly,

$$\hat{G}_c(t) = \hat{\mu}(g_c, H_1) \sum_{i=1}^m \hat{H}_1^{-1}(y_i) \hat{q}_i I(y_i \leq t),$$

$$\text{where } \hat{\mu}(g_c, H_1) = \left\{ \sum_{i=1}^m \frac{\hat{q}_i}{\hat{H}_1(y_i)} \right\}^{-1}.$$

The cumulative incidence in each treatment arm can be estimated nonparametrically by the method of Maller and Zhou (1992),

$$\hat{\pi}_0 = \tilde{F}_{n_0}(x_{(n_0)}), \quad \hat{\pi}_1 = \tilde{G}_{n_1}(x_{(n_1)}),$$

where $1 - \tilde{F}_{n_0}$ and $1 - \tilde{G}_{n_1}$ are the conventional Kaplan-Meier estimators of the placebo and TAM, respectively, and $x_{(n_0)}$ and $x_{(n_1)}$ are the last observed failure times or censoring times in the placebo and TAM arms, respectively. Without using the assumed functional relationship between f_c and g_c , we separately estimate these two distributions nonparametrically using data from each of the two arms. Given the above estimated rates, the conditional cdfs F_c and G_c can be derived by

$$\tilde{F}_c(t) = \hat{\pi}_0^{-1} \tilde{F}_{n_0}(t), \quad \tilde{G}_c(t) = \hat{\pi}_1^{-1} \tilde{G}_{n_1}(t).$$

These nonparametric estimators will be utilized for validating the semiparametric model assumption (1.1) in the next section.

3 INFERENCE

3.1 Goodness-of-Fit Test for Model Checking

Due to the parametric component imposed on the ratio of two density functions in model (1.1), model-checking procedures are needed to examine the adequacy of the link function. There is limited literature for goodness-of-fit tests in this type of setting. By generalizing the bootstrap resampling approaches of Qin and Zhang (1997), we propose a test statistic based on the area between two survival curves for right-censored data.

To test stochastic ordering of two survival distributions, Pepe and Fleming (1989) developed a class of statistics based on the integrated weighted difference in Kaplan-Meier estimators. For the purpose of general model checking, it is more appropriate to use a quadratic form of the cumulative difference between the two curves to reveal any departure from the model assumption instead of ordering. Formally, under the assumptions $\lim_{n \rightarrow \infty} n_0/n > 0$ and τ being bounded, we define the test statistic as

$$\Delta_n = \int_0^\tau w(t) [\widehat{F}_c(t) - \tilde{F}_c(t)]^2 dt, \quad (3.1)$$

where $\tau = \min(x_{(n_0)}, x_{(n_1)})$, and $w(t)$ is a positive weight function, which may be used to improve the efficiency of the test. The choice of the weight functions has been discussed by Pepe and Fleming (1989), and relies on stability considerations. Misspecification of the model can be identified by a large area between $F_c(t)$ and $F_c^\wedge(t)$. Recall that $F_c^\wedge(t)$ is estimated based on data from both placebo and TAM arms under assumption (1.1), and $F_c^\wedge(t)$ is estimated nonparametrically from data in the placebo arm only.

The asymptotic distribution of Δ_n is more complicated than the Pepe-Fleming type of test statistics, since the overlapping data are used for estimating $F_c^\wedge(t)$ and $F_c(t)$. However, its theoretical properties can be derived in a manner similar to that given by Qin and Zhang (1997). Since it is not the primary focus of this paper, we use a bootstrap method to find the critical values of the test statistic, Δ_n . Let the time of disease diagnosis $y_1^*, y_2^*, \dots, y_{m_0}^*$ be drawn independently with replacement from $dF_1^\wedge$ for the placebo arm and let $y_{m_0+1}^*, y_{m_0+2}^*, \dots, y_m^*$ be drawn independently from $dG_1^\wedge$ for the treatment arm. The censoring times for each treatment arm are drawn independently with replacement from the raw data within each arm. Based on the bootstrap samples of all failure and censoring times, we can estimate $\widehat{F}_c^*(t)$ under the semiparametric model assumption (1.1), and nonparametric $\widehat{F}_c^*(t)$. The bootstrap test statistic $\Delta_n^* = \int_0^\tau w^*(t) [\widehat{F}_c^*(t) - \tilde{F}_c^*(t)]^2 dt$ can be calculated. With large repetitions of the re-sampling, we can estimate the critical values of the test statistic Δ_n .

In a typical cancer prevention trial, a majority of the participants will never experience the event of interest, and will be censored at end of the follow-up. For example, in the BCPT trial more than 95% of the participants were censored during the study follow-up. Re-sampling the censored times from this large, raw data set is computationally intensive and may not add much precision to the estimation of the critical values of the test statistic. Therefore, we choose to re-sample the failure times only, using the same procedures described above. Similarly, the bootstrap samples, $y_1^*, y_2^*, \dots, y_m^*$ for both arms are used to derive the estimator $F_1^*(t)$. At re-sampled event time y_j^* , we use the estimator of $\widehat{\psi}(y_j^*)$, where $\widehat{\psi}$ was estimated from the raw data. This alternative method has the computational advantage of directly checking the model assumption on f_1 and g_1 . In this case, $\tilde{F}_1^*(t) = m_0^{-1} \sum_{i=1}^{m_0} I(y_i^* \leq t)$ is an empirical distribution function based on $y_1^*, \dots, y_{m_0}^*$. The corresponding bootstrap version of the test statistic is

$$\Delta_{1n}^* = \int_0^\tau w^*(t) [\widehat{F}_1^*(t) - \tilde{F}_1^*(t)]^2 dt.$$

3.2 Inference on Treatment Effects

When the model assumption (1.1) is validated, the next issue is to assess the effect of treatment on the time to the disease diagnosis and on the disease incidence. For the time to the diagnosis, we are interested in testing $H_{0A} : \alpha = \alpha_0, \beta = \beta_0$ in model (1.1). When $\beta_0 = \alpha_0 = 0$, the two density functions are equal, $f_c(t) = g_c(t)$. When $\beta = 1$ for a given $h(t) = \log(t)$, the model (1.1) reduces to a length-biased sampling problem.

When the censoring distributions of the two arms are equal, $\psi = 0$, we can use following conditional likelihood ratio test statistic to test $\beta = \beta_0$ while letting α be a nuisance parameter,

$$R_n = 2(\ell_c(\hat{\alpha}, \hat{\beta}) - \ell(\tilde{\alpha}_0, \beta_0)),$$

where $\tilde{\alpha}_0$ is the conditional MLE given $\beta = \beta_0$. R_n converges to a chi-square distribution with a degree of freedom of 1. When $\psi(t) = 0$, $\beta = 0$ implies that $\alpha = 0$ and the two density functions are equal. Then the likelihood ratio statistic, $R_n = 2(\ell_c(\hat{\alpha}, \hat{\beta}) - \ell(0, 0))$, converges to a chi-square distribution with a degree of freedom of 1, as well. When $\psi(\neq 0)$ is estimated by its consistent estimator, the asymptotic distribution of R_n under H_{0A} is a weighted sum of the chi-square variables (Liang and Self, 1996). The derivation of its asymptotic distribution is tedious and not the focus here, thus we omit it. However, we can still obtain the valid critical value for the test statistic using bootstrap methods.

To conduct a formal inference on the difference of the cumulative incidences of the two arms, we test the null hypothesis: $H_{0\pi} : \pi_0 = \pi_1$. It is clear that we cannot rely on the conditional likelihood to test $H_{0\pi}$, because the incidence in each arm is canceled out. Instead, we propose a test statistic based on the weak convergence of $\hat{\pi}_0 - \hat{\pi}_1$ from Theorem 2 of Maller and Zhou (1992). Specifically,

$$\sqrt{n_0}(\hat{\pi}_0 - \pi_0) \rightarrow N(0, (1 - \pi_0)^2 \nu_0), \quad \sqrt{n_1}(\hat{\pi}_1 - \pi_1) \rightarrow N(0, (1 - \pi_1)^2 \nu_1),$$

where ν_0 can be consistently estimated by

$$\hat{\nu}_0 = \int_0^{X(n_0)} \frac{n_0 I(\Delta N(t) < Y_{n_0}(t)) dN_0(t)}{Y_{n_0}(t)(Y_{n_0}(t) - \Delta N(t))},$$

$Y_{n_0}(t) = \sum_{i=1}^{n_0} I(x_i \geq t)$, $\Delta N(t)$ is the jump in $N(\cdot)$ at time $t > 0$, and $N_0(t) = \sum_{i=1}^{n_0} I(x_i \leq t, \delta_i = 1)$. Similarly, ν_1 can be estimated from the data in the TAM arm. The statistic for testing $H_{0\pi}$ can be constructed as $T_n = \sqrt{n}(\hat{\pi}_0 - \hat{\pi}_1)$. Under $H_{0\pi}$:

$$T_n = \sqrt{1+\eta} \sqrt{n_0}(\hat{\pi}_0 - \pi_0) - \sqrt{1+1/\eta} \sqrt{n_1}(\hat{\pi}_1 - \pi_1),$$

where $\eta = n_1/n_0$. Under the null hypothesis of $H_{0\pi} : \frac{T_n}{\{(1+\eta)\hat{\nu}_0 + (1+1/\eta)\hat{\nu}_1\}(1-\hat{\pi})^2}$ is asymptotically standard normal, where $\hat{\pi}$ is estimated using data from both arms. Reject $H_{0\pi}$, if $|T_n| > z_{\alpha/2}$ for a two-sided test at level α .

4 Simulations

In the simulation study, we assess the point estimators and their standard deviations of α and β under model (1.1) with $h(t) = t$, and the nonparametric estimators for π_0 and π_1 . We also examine the size and power properties of the test statistic for goodness-of-fit to model (1.1), the conditional likelihood ratio test for α and β , and the asymptotic normal test for comparing the incidence rates between the two arms. We list the results of simulations repeated 5000 times in Table 1 and 1000 times in Table 2. The sample size used in each arm is 200 or 300.

Assume disease incidences: $\pi_0 = 0.2$ or 0.3 and $\pi_1 = 0.2$. Among incident cases, we generated the times to diagnosis of the disease in the placebo arm using an exponential distribution with mean ω for f_c , and generated the censoring times for both arms from the same uniform distribution. Because of the low incidence of the disease, the overall censoring percentages are

high, and vary from 74% to 90% under different scenarios. We related g_c and f_c by the following equation

$$g_c(t) = t^{-\gamma} \exp(\alpha + \beta t + \gamma t^2) f_c(t). \quad (4.1)$$

When $\gamma = 0$, the above model reduces to (1.1) with $h(t) = t$. Thus, g_c is an exponential pdf with mean v , $\alpha = \log(\omega) - \log(v)$, and $\beta = 1/\omega - 1/v$. Let $\omega = 1.5$ and $v = 1.5, 2.5$, or 3 , which correspond to various combinations of α and β values in Table 2.

Under the alternative hypothesis to model (1.1), we let $\gamma = -1$ to simplify the simulations, so that g_c is a pdf of Weibull variables. Specifically,

$$g_c(t) = \frac{\exp\{\alpha + (\beta - 1/\omega)^2/4\}}{\omega} t \exp\{-(t - [\beta - 1/\omega])^2\}.$$

Letting the location parameter be zero in Weibull, the parameters in model (4.1) have $\alpha = \log(2\omega)$, and $\beta = 1/\omega$.

The estimated parameters α , β , π_0 and π_1 under the exponential tilt model (1.1) are summarized in Table 1. The size and power obtained for testing the null hypotheses H_{0A} and $H_{0\pi}$ under each scenario are also shown in this table. The first horizontal panel in Table 1 shows that no serious bias is seen in the point estimators of α and β from the conditional likelihood, and the estimators for incidence rates under H_{0A} and $H_{0\pi}$. The variations of the estimators decrease with sample size. The significance levels of the two test statistics, R_n and T_n , are maintained at the nominal level of 0.05. The second panel of Table 1 shows the results when $f_c = g_c$ (i.e. under H_{0A}) but with different incidence rates for the two arms. Similar results are observed. The power for testing $H_{0\pi}$ increased with the sample size.

The next four panels present scenarios with a mean time to event in the placebo arm of $\omega = 1.5$, and $v = 2.5$ (or $v = 3$) for the TAM treatment arm (corresponding to $\alpha = -0.4, \beta = 0.267$ or $\alpha = -0.519, \beta = 0.333$). Because of the small number of events and the heavy censoring in the treatment arm, the type I error rate for testing $\pi_0 = \pi_1$ was somewhat inflated. The incidence for the treatment arm is also slightly underestimated in this case. When incidence rates are different, the power of the test statistic T_n increases with the sample size, and the power of test statistic R_n increases with the number of the events.

The upper panels (from the first to the fourth) in Table 2 list the size of the test statistic Δ_{1n} under the model (1.1) with $h(t) = t$, and the size and power of T_n for testing $H_{0\pi}$ with different combinations of α , β and incidence rates. For significance levels 0.10, 0.05 and 0.025, we approximated the corresponding critical values of Δ_{1n} by using the appropriate values in the 500 ordered bootstrap replications of Δ_{1n}^* . The lower panels in Table 2 show the power of Δ_{1n} under the alternative model (4.1) with $\gamma = -1$, and the size or power of T_n . The power of Δ_{1n} increases with a higher incidence and/or a larger sample size. We also compared the performance of Δ_{1n}^* with Δ_n^* , which required us to resample the whole data set, including the censoring times. Both test statistics have maintained the specified significance levels, but Δ_n is slightly more conservative than Δ_{1n} under the null hypothesis. As expected, the powers of Δ_{1n} are greater than those of Δ_n . Note that test statistic T_n is valid even under a misspecified model (1.1), because the asymptotic property of T_n does not rely on model (1.1).

5 Application to the NSABP Study

We return to the NSABP-BCPT trial that motivated this study. We use the same data set, including 6599 participants on placebo and 6576 on TAM, as reported in the NSABP (P-1) analyses (Fisher et al, 1998). The median age at diagnosis was 55 and 53 years for the placebo and TAM groups, respectively. We will focus on the time from randomization to the diagnosis of breast cancer and the incidence. All trial participants received screening examinations at the same frequency, as described in the protocol.

5.1 Overall Analyses

Applying the proposed methods, we first checked the validity of the model (1.1) with the identity link for h . By using the bootstrap method and re-sampling data from all the participants, we obtained a P -value of 0.742 for the proposed test (Δ_n) with $w(t) = 1$. Using the alternative re-sampling method and focusing on the observed cases only, the test statistic (Δ_{1n}) leads to a P -value of 0.320. Under model (1.1), the unknown parameters α and β and their standard deviations were estimated and listed in Table 3. In Figure 1, the model-based survival curve for the placebo arm, $(1 - F^{\wedge}_c(t))$, was plotted against the nonparametric conditional survival curve, $(1 - F^{\wedge}_c(t))$, based on data from the placebo arm only. The empirical results, including the estimated survival curves, indicate that model (1.1) fits the data well.

The censoring distributions of the two arms show little difference (P -value = 0.663), thus ψ approximates to zero in (2.2). The P -value for the statistic, T_n , which tests $\alpha = \beta = 0$ based on the conditional likelihood ratio test is 0.598. In Figure 2, we present the model-based survival function for each arm conditional on the diagnosed invasive breast cancers, $1 - F^{\wedge}_1$ and $1 - G^{\wedge}_1$ (Figure 2A), and conditional on all (expected) invasive cases, $1 - F^{\wedge}_c$ and $1 - G^{\wedge}_c$ (Figure 2B). Both plots imply that the survival curves for the two arms are fairly close to each other.

The estimated breast cancer incidence in the placebo and TAM arms is 0.044 and 0.022, respectively. The corresponding P -value for testing $\pi_0 = \pi_1$ is 0.0012. Based on the proposed methods, we conclude that the treatment effect of TAM is substantial in the reduction of invasive breast cancer events. However, among women who were diagnosed with breast cancer during the study follow-up, there was little difference in terms of time to the disease diagnosis.

5.2 Subset Analysis by Estrogen Receptor Status

To investigate the treatment effects of TAM within sub-cohorts of patients as determined by the ER status of their tumors, we performed analyses for each sub-cohort by ER status. Among the diagnosed invasive breast cancers, there were a total of 174 patients with ER-positive tumors, and 69 patients with ER-negative tumors. Among the ER-negative cohort, the median ages at diagnosis were similar being 52.5 years and 52.7 years for the placebo and TAM arms, respectively. Among the ER-positive cohort, the median age at diagnosis among the placebo arm was about two years older than that for the TAM arm being 55.8 and 53.5, respectively. Among ER-positive breast cancers, the median time (from randomization) to diagnosis is 27 months for placebo and 30 months for TAM. Among invasive ER-negative breast cancers, the median time to diagnosis is 30 months for placebo and 23 months for TAM.

The tests for checking the model fitting with $h(t) = t$ for ER-positive tumors lead to P -values of 0.635 and 0.624 for Δ_n and Δ_{1n} . The empirical plot in Figure 3 confirms the analytical results. The conditional likelihood ratio test, R_n , results in a P -value of 0.48, which shows no significant effect for TAM on the time to disease diagnosis among patients with ER-positive tumors, as shown in Figure 4. The estimated disease incidence for patients receiving TAM and placebo is 0.013 and 0.035, respectively. The corresponding statistic for testing $H_{0\pi}$ has a P -value of 0.001. These results are consistent with the overall analysis of the cohort, but show more

substantial benefits in the reduction of incidence for women with ER-positive tumors compared to the pooled cohorts.

Similar to the findings for the ER-positive cohort, the model (1·1) with the identity link h has a good fit with the data for the ER-negative cohort, which is shown in Figure 5 (P -value = 0.94). In contrast to the findings for the ER-positive cohort, the conditional likelihood ratio test under the exponential tilt model for the ER-negative cohort indicates a statistically significant effect of TAM on the time to diagnosis (P -value=0.037). Compared to patients in the placebo arm, time to the diagnosis of breast cancer was shifted (earlier) for the patients in the TAM arm. This is also shown in Figure 6B that the median time to the disease diagnosis is 24 months and 36 months for the TAM and placebo arm, respectively. Note that 58% of the ER-negative tumors in the tamoxifen group were diagnosed in the first two years of the study, compared with 29% in the placebo group. The estimated disease incidences for TAM and placebo are about the same: 0.007 for placebo and 0.0074 for TAM, with a P -value of 0.828 for test statistic T_n . As suggested by one reviewer, we conducted a simulation to assess the power of test statistic, T_n , to detect a two-fold decrease (or increase) in incidence in the TAM arm compared to that in the placebo with $\pi_0 = .007$. In a setting similar to that of the BCPT trial, the power of the test would be adequate to detect a two-fold decrease (or increase) in incidence with 78.4% (or 97.5%) at the significance level of 0.05.

6 Discussion

We have investigated a class of semiparametric models other than the commonly used survival analysis models within the cure rate framework. The class of flexible models that includes the exponential tilt and generalized length-biased models has not been adequately explored for censored survival data. We have generalized the method of Qin and Zhang (1997) for case-control studies with immediately observed data to randomized clinical trials with censored survival data. The proposed semiparametric approaches can then be used to estimate and test two treatment effects on disease incidence and on time to the disease diagnosis separately. The estimation and inference procedures are relatively straightforward without requiring the use of the EM algorithm, and can be easily implemented. The S-plus (or R) functions we used to fit the model and test the multiple hypotheses can be obtained on request.

Based on the NSABP-BCPT data, Radmacher and Simon (2000) used tumor growth models to explore the role of tamoxifen in the prevention of breast cancer. Their results suggest that tamoxifen reduce cancer incidence in the BCPT through both treating occult disease and preventing new tumor formation and growth. By taking a different perspective and approach to that of Radmacher and Simon, we model both the cumulated incidence of the disease and the time to diagnosis of the disease for patients, so that we can discriminate the effects of the treatment on the two factors. Our findings on the disease incidence agree with that of most of the literature indicating that tamoxifen can prevent new tumor formation, which is reflected in the reduction of the ER-positive disease incidence. Although there is no evidence that tamoxifen reduces (or increases) the incidence of breast cancer for ER-negative tumors, our analyses indicated that TAM reduced the time to diagnosis of breast cancer among patients with ER-negative tumors. Indeed, the earlier diagnosis may be a result of more “diagnosable” tumors or a result of faster growing tumors associated with the use of TAM. If TAM accelerates the growth of tumors in ER-negative patients, we would expect to see more cases in the TAM group than in the placebo group, which seems not the case for this study. Further insight into the possible mechanism for this can be gained by comparing the distributions of tumor size at diagnosis among ER-negative tumors. As shown in Table 4, we see that the distribution of tumor size at the diagnosis in the TAM arm is shifted to a smaller size than is that in the placebo arm. In the TAM arm, 39.5% of the tumors were under 1 cm in size, compared to 16.1% in the placebo group (the exact P -value of the Wilcoxon-Mann-Whitney test is 0.045). This

observation inclines to support the conjecture that the use of TAM expedites the diagnosis of breast cancer for ER negative tumors. Some recent studies found that hormone-based tamoxifen treatment renders occult tumor tissue less dense, which may lead to more readily diagnosable diseased cases (Atkinson et al, 1999; Brisson et al, 2000; Chow et al, 2000).

The findings from our study have the potential to impact research in the effect of tamoxifen on the natural history of breast cancer and on both primary and secondary prevention of the disease. We also acknowledge our limitation to give a definitive interpretation of why TAM affects the time to diagnosis of ER-negative tumors. The biologic mechanism for the shift in diagnosis caused by TAM on ER-negative tumors should be further investigated in clinical studies, so that we may better understand the role of TAM in the prevention of breast cancer.

Our models and methods can be extended in several directions. Breast cancer risk is age-dependent. The current model provides an average treatment effect on the time to diagnosis of breast cancer for this cohort of subjects. More general types of semiparametric models may be constructed to take covariates other than the treatment indicator (such as ER status and age at study entry) into consideration. These baseline covariates can be included in separately modeling the disease incidence and in time to the event. However, the inference and model validation procedures are considerably complicated, and beyond the scope of this study.

Since conditional inference is often less efficient than an inference based on the full likelihood, it would be interesting to assess the loss of efficiency due to conditioning. However, it can be difficult to obtain the maximum likelihood estimators under the semi-parametric model in the current setup, based on the full likelihood. Intuitively, the loss of efficiency in the conditional inference should be limited in the current context. Since (α, β) are the regression parameters describing the treatment effect on the time to diagnosis among only the diseased cases, the conditional inference based on all observed diseased cases should contribute most of the information on this part of the inference. In a typical cancer prevention trial, a majority of trial participants will not have the disease of interest (either within the follow-up duration or in their lifetime). Using the full likelihood would be computationally inefficient because most of the censored subjects would not be diagnosed with the disease even if they had not been censored.

Acknowledgments

This research was partially supported by National Institute of Health grants R01-CA079466, U10-CA37377 and U10-CA69974. The authors wish to thank the editor, the associate editor, and two referees for their constructive comments.

REFERENCES

- Anderson JA. Multivariate logistic compounds. *Biometrika* 1979;66:17–26.
- Atkinson C, Warren R, Bingham SA, Day NE. Mammographic patterns as a predictive biomarker of breast cancer risk: effect of tamoxifen. *Cancer Epidemiol Biomarkers Prev* 1999;8:863–866. [PubMed: 10548313]
- Brisson J, Brisson B, Cote G, Maunsell E, Berube S, Robert J. Tamoxifen and mammographic breast densities. *Cancer Epidemiol Biomarkers Prev* 2000;9:911–915. [PubMed: 11008908]
- Broet P, De Rycke Y, Tubert-Bitter P, Lellouch J, Asselain B, Moreau T. A Semiparametric Approach for the Two-sample Comparison of Survival Times with Long-term Survivors. *Biometrics* 2001;57:844–852. [PubMed: 11550936]
- Chen MH, Ibrahim JG, Sinha D. A New Bayesian Model for Survival Data With a Surviving Fraction. *Journal of the American Statistical Association* 1999;94:909–919.
- Cheng SC, Wei LJ, Ying Z. Analysis of transformation models with censored data. *Biometrika* 1995;82:835–845.
- Chow CK, Venzon D, Jones EC, Premkumar A, O'Shaughnessy J, Zujewski J. Effect of Tamoxifen on Mammographic Density. *Cancer Epidemiology and Biomarkers Prevention* 2000;9:917–921.

- Cox, DR. Some sampling problems in technology. In: Johnson, NL.; Smith, H., Jr, editors. *New Development in Survey Sampling*. New York: Wiley Interscience; 1969. p. 506-527.
- Cuzick J, Powles T, Veronesi U, Forbes J, Edwards R, Ashley S, Boyle P. Overview of the main outcomes in breast-cancer prevention trials. *Lancet* 2003;361:296–300. [PubMed: 12559863]
- Early Breast Cancer Trialists Collaborative Group. Effects of Chemotherapy and Hormonal Therapy for Early Breast Cancer on Recurrence and 15-year Survival: An Overview of the Randomised Trials. *Lancet* 2005;365:1687–1717. [PubMed: 15894097]
- Farewell V. The Use of Mixture Models for the Analysis of Survival Data With Long-term Survivors. *Biometrics* 1982;38:1041–1046. [PubMed: 7168793]
- Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, Vogel V, Robidoux A, Dimitrov N, Atkins J, Daly M, Wieand S, Tan-Chiu E, Ford L, Wolmark N. Tamoxifen for Prevention of Breast Cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *Journal of National Cancer Institute* 1998;90:1371–1388.
- Fisher B, Dignam J, Wolmark N, Wickerham DL, Fisher ER, Mamounas E, Smith R, Begovic M, Dimitrov NV, Margolese RG, Kardinal CG, Kavanah MT, Fehrenbacher L, Oishi RH. Tamoxifen in Treatment of Intraductal Breast Cancer: National Surgical Adjuvant Breast and Bowel Project B-24 Randomised Controlled Trial. *Lancet* 1999;353:1993–2000. [PubMed: 10376613]
- Ibrahim, JG.; Chen, MH.; Sinha, D. *Bayesian Survival Analysis*. New York: Springer-Verlag; 2001a.
- Ibrahim JG, Chen MH, Sinha D. Bayesian Semi-Parametric Models for Survival Data With a Cure Fraction. *Biometrics* 2001b;57:383–388. [PubMed: 11414560]
- Kaufmann M, Graf E, Jonat W, Eiermann W, Geberth M, Albert U, Gademann G, Conrad B, Stahl K, Minckwitz G, Schumacher M. The German Adjuvant Breast Cancer Group. Tamoxifen Versus Control after Adjuvant, Risk-adapted Chemotherapy in Postmenopausal, Receptor-negative Patients with Breast Cancer: A randomized Trial (GABG-IV D-93). *Journal of Clinical Oncology* 2006;23:7842–7848. [PubMed: 16258087]
- Kay R, Little S. Transformations of the Explanatory Variables in the Logistic Regression Model for Binary Data. *Biometrika* 1987;74:495–501.
- Liang KY, Steven SG. On the asymptotic behaviour of the pseudo-likelihood ratio test statistic. *Journal of the Royal Statistical Society, Series B, Methodological* 1996;58:785–796.
- Maller RA, Zhou S. Estimating the Proportion of Immunes in a Censored Sample. *Biometrika* 1992;79:731–739.
- Qin J, Leung D. A Semiparametric Two-component Compound Mixture Model and Its Application to Estimating Malaria Attributable Fractions. *Biometrics* 2005;61:456–464. [PubMed: 16011692]
- Qin J, Berwick M, Ashbolt R, Dwyer T. Quantifying the Change of Melanoma Incidence by Breslow Thickness. *Biometrics* 2002;58:665–670. [PubMed: 12230002]
- Qin J, Zhang B. A Goodness-of-fit Test for Logistic Regression Models Based on Case-control Data. *Biometrika* 1997;84:609–618.
- Radmacher MD, Simon R. Estimation of Tamoxifen's Efficacy for Preventing the Formation and Growth of Breast Tumors. *Journal of National Cancer Institute* 2000;92:48–53.
- Simon R. Length Biased Sampling in Etiologic Studies. *American Journal of Epidemiology* 1980;111:444–452. [PubMed: 7377187]
- Sy JP, Taylor JMG. Estimation in a Cox Proportional Hazards Cure Model. *Biometrics* 2002;56:227–236. [PubMed: 10783800]
- Tsodikov A. A Proportional Hazards Model Taking Account of Longterm Survivors. *Biometrics* 1998;54:1508–1516. [PubMed: 9883549]
- Tsodikov A. Semi-parametric Models of Long-and Short-term Survival: An Application to the Analysis of Breast Cancer Survival in Utah by Age and Stage. *Statistics in Medicine* 2002;21:895–920. [PubMed: 11870824]
- Tsodikov A, Ibrahim IG, Yakovlev AY. Estimating Cure Rates from Survival Data: An Alternative to Two-component Mixture Models. *Journal of the American Statistical Association* 2003;98:1063–1078.
- Vardi Y. Empirical Distribution in Selection Bias Models. *Annals of Statistics* 1985;13:178–203.
- Wang MC. Hazards Regression Analysis for Length-biased Data. *Biometrika* 1996;83:343–354.

- White IR, Thompson SG. Choice of Test for Comparing Two Groups, with Particular Application to Skewed Outcomes. *Statistics in Medicine* 2003;22:1205–1215. [PubMed: 12687651]
- Zelen M, Feinleib M. On the Theory of Screening for Chronic Diseases. *Biometrika* 1969;56:601–614.

Figure 1. Survival Distributions of Placebo Arm Given All Incidence Cases for Model Checking: Estimated under Exponential-tilt Model Assumption using Data from Both Arms(—), and Nonparametric Estimator using Data from Placebo Arm Only (- - -)

Figure 2. Estimated Survival Distributions under Exponential-tilt Model for Each Treatment Arm (A) Conditional on Diagnosed Cases and (B) Conditional on Prevalent Cases

Figure 3. Survival Distributions of Placebo Arm Given All Incidence Cases for Model Checking in ER-positive Tumors: Estimated under Exponential-tilt Assumption using Data from Both Arms (—), and Nonparametric Estimator using Data from Placebo Arm Only (---)

Figure 4. Estimated Survival Distributions under Exponential-tilt Model for Each Treatment Arm Among Women with ER-Positive Tumors (A) Conditional on Diagnosed Cases and (B) Conditional on Prevalent Cases

Figure 5. Survival Distributions of Placebo Arm Given All Incidence Cases for Model Checking in ER-negative tumors: Estimated under Exponential-tilt Model Assumption using Data from Both Arms(-), and Nonparametric Estimator using Data from Placebo Arm Only (- -)

Figure 6. Estimated Survival Distributions under Exponential-tilt Model for Each Treatment Arm Among Women with ER-negative Tumors (A) Conditional on Diagnosed Cases and (B) Conditional on Prevalent Cases

Table 1

Parameter Estimation and its SD, Size and Power by Test Statistics under Different Scenarios with Exponential-tilt Model

Parameter	Est. (SD)	Tests (size/power)	Est. (SD)	Tests (size/power)
	$n_0 = n_1 = 200$		$n_0 = n_1 = 300$	
$\alpha = 0$	-.003(.284)		-.004(.226)	
$\beta = 0$	-.002(.219)	.053	.003(.175)	.057
$\pi_0 = .200$	.196(.029)		.197(.025)	
$\pi_1 = .200$	.195(.030)	.056	.196(.027)	.049
$\alpha = 0$	.001(.253)		.001(.207)	
$\beta = 0$	-.004(.197)	.047	-.003(.161)	.050
$\pi_0 = .300$	.295(.034)		.294(.029)	
$\pi_1 = .200$	.196(.029)	.614	.198(.024)	.774
$\alpha = -.400$	-.419(.317)		-.411(.254)	
$\beta = .267$	.281(.214)	.294	.274(.170)	.418
$\pi_0 = .200$	.196(.030)		.196(.028)	
$\pi_1 = .200$	.182(.034)	.072	.184(.027)	.065
$\alpha = -.400$	-.418(.292)		-.405(.226)	
$\beta = .267$	.277(.192)	.338	.269(.148)	.465
$\pi_0 = .300$	.295(.033)		.293(.028)	
$\pi_1 = .200$	.182(.032)	.721	.183(.025)	.855
$\alpha = -.519$	-.544(.344)		-.540(.270)	
$\beta = .333$	.351(.222)	.414	.346(.174)	.581
$\pi_0 = .200$	.197(.031)		.196(.028)	
$\pi_1 = .200$	.172(.033)	.103	.184(.027)	.110
$\alpha = -.519$	-.542(.312)		-.534(.269)	
$\beta = .333$	.346(.194)	.485	.342(.152)	.646
$\pi_0 = .300$	.295(.035)		.296(.030)	
$\pi_1 = .200$	.173(.034)	.773	.174(.027)	.891

Table 2
Size and Power of Test Statistics for Model Checking and Comparing Incidence under Different Scenarios

Parameters	Levels	Δ_{In}	T_n	$n_0 = n_1 = 200$	$n_0 = n_1 = 300$	T_n
Under the null model (1·1), $h(t) = t$						
$\pi_0 = \pi_1 = 0.2, \alpha = \beta = 0$	.100	.074	.104	$\gamma = 0$	.081	.098
	.050	.033	.056		.035	.049
	.025	.019	.029		.023	.026
$\pi_0 = 0.3, \pi_1 = 0.2, \alpha = \beta = 0$	.100	.057	.725	.084	.856	
	.050	.027	.614	.047	.774	
	.025	.012	.499	.027	.678	
$\pi_0 = \pi_1 = 0.2, \alpha = -.40, \beta = .267$	.100	.076	.120	.085	.135	
	.050	.041	.072	.032	.065	
	.025	.018	.032	.019	.035	
$\pi_0 = 0.3, \pi_1 = 0.2, \alpha = -.40, \beta = .267$	.100	.069	.805	.074	.914	
	.050	.039	.721	.036	.855	
	.025	.024	.629	.020	.793	
Under the alternative (4·1)						
$\pi_0 = \pi_1 = 0.2, \omega = 2$	.100	.638	.110	$\gamma = -1$	.903	.118
	.050	.507	.059		.843	.051
	.025	.382	.031		.765	.028
$\pi_0 = \pi_1 = 0.2, \omega = 1.5$	.100	.759	.108	.955	.110	
	.050	.637	.056	.905	.055	
	.025	.532	.029	.844	.028	
$\pi_0 = 0.3, \pi_1 = 0.2, \omega = 2$	.100	.780	.600	.973	.764	
	.050	.690	.468	.944	.648	

Parameters	Levels	Δ_{1n} $n_0 = n_1 = 200$	T_n	Δ_{1n} $n_0 = n_1 = 300$	T_n
$\pi_0 = 0.3, \pi_1 = 0.2, \omega = 1.5$	.025	.573	.354	.905	.548
	.100	.880	.698	.987	.837
	.050	.800	.566	.957	.750
	.025	.714	.446	.925	.644

Table 3

Estimated Parameters (SD) in the Proposed Model and P-values of Test Statistics for the NSABP-BPCT Trial

	Overall	ER+	ER–
$\hat{\pi}_1$ (SD)	0.022 (.004)	0.013 (.004)	0.0074 (.001)
$\hat{\pi}_0$ (SD)	0.044 (.005)	0.035 (.005)	0.0070 (.001)
R_n , P-value	0.001	<0.001	0.828
$\hat{\alpha}$ (SD)	0.124 (0.233)	−0.218 (.324)	1.120 (.532)
$\hat{\beta}$ (SD)	−0.001 (0.008)	0.007 (.010)	−0.035 (.018)
T_n , P-value	0.598	0.480	0.037
Δ_{1n} , P-value	0.320	0.624	0.940

Table 4

Tumor Size at the Diagnosis by Treatment among ER Negative Tumors

Tumor size (cm)	Placebo		Tamoxifen	
	N	%	N	%
≤ 1.0	5	16.1	15	39.5
1.1 – 2.0	14	45.2	13	34.2
2.1–3.0	6	19.4	7	18.4
≥3.1	5	16.1	3	7.9
Unknown	1	3.2	0	0