Contemporary Clinical Trials 59 (2017) 30–37

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Cure modeling in real-time prediction: How much does it help?☆

Gui-shuang Ying[a,b], Qiang Zhang[c], Yu Lan[d], Yimei Li[b], Daniel F. Heitjan[d,e,*]

[a] Department of Ophthalmology, University of Pennsylvania, Philadelphia, PA 19104, USA
[b] Department of Biostatistics & Epidemiology, University of Pennsylvania, Philadelphia, PA 19104, USA
[c] NRG Oncology Statistics & Data Management Center, Philadelphia, PA 19103, USA
[d] Department of Statistical Science, Southern Methodist University, Dallas, TX 75275, USA
[e] Department of Clinical Sciences, UT Southwestern Medical Center, Dallas, TX 75390, USA

ARTICLE INFO

Keywords:
Bayesian bootstrap
Enrollment model
Event-based trial
Interim analysis
Weibull distribution

ABSTRACT

Various parametric and nonparametric modeling approaches exist for real-time prediction in time-to-event clinical trials. Recently, Chen (2016 BMC Biomedical Research Methodology **16**) proposed a prediction method based on parametric cure-mixture modeling, intending to cover those situations where it appears that a non-negligible fraction of subjects is cured. In this article we apply a Weibull cure-mixture model to create predictions, demonstrating the approach in RTOG 0129, a randomized trial in head-and-neck cancer. We compare the ultimate realized data in RTOG 0129 to interim predictions from a Weibull cure-mixture model, a standard Weibull model without a cure component, and a nonparametric model based on the Bayesian bootstrap. The standard Weibull model predicted that events would occur earlier than the Weibull cure-mixture model, but the difference was unremarkable until late in the trial when evidence for a cure became clear. Nonparametric predictions often gave undefined predictions or infinite prediction intervals, particularly at early stages of the trial. Simulations suggest that cure modeling can yield better-calibrated prediction intervals when there is a cured component, or the appearance of a cured component, but at a substantial cost in the average width of the intervals.

1. Introduction

Many clinical trials with time-to-event outcomes schedule interim and final analyses to take place on the occurrence of a pre-specified number of events. For example, a cancer trial could be designed to have 80% statistical power with 300 deaths, with planned interim analyses after the 100th and 200th deaths, and a final analysis when the 300th death occurs [1]. Because the times of occurrence of these landmark events are random, it is desirable to have a tool for predicting them as an aid to logistical planning.

We have developed a range of models for making such predictions [2–6] and demonstrated their utility in a motivating clinical trial [7]. These models assume that every participant is susceptible and will eventually experience the event if follow-up time is sufficiently long [8,9]. This assumption may not hold in diseases where there is a possibility of cure — for example, in many childhood cancers and some adult cancers such as leukemia [10,11], colon cancer [12], and head-and-neck cancer [13]. Failure to accommodate the possibility of cure could in principle lead to bias, because a fraction of surviving patients would be predicted to experience events to which they are effectively no longer susceptible.

Recognizing this gap in the literature, Chen [14] recently proposed the use of cure models [15] in prediction. His method involves, at each prediction time, selecting the best-fitting from among a menu of cure-mixture models (exponential, Weibull, log-logistic and lognormal), and using it to create predicted values for all as-yet unenrolled subjects and all enrolled but censored subjects. He demonstrated the application of his method in a two-arm cancer immunotherapy trial. At every prediction time, goodness-of-fit statistics selected the Weibull cure-mixture model. Initial predictions of the time of the planned final analysis (to be conducted at the 416th death) were as much as 20 months early, but subsequent predictions were closer to the target. Only at the final two (of six) predictions did the empirical survival curve reveal the characteristic shape of the cure model.

Chen presented no evaluation of how his method would perform in repeated samples, either as an estimation or a prediction procedure. Although parametric cure-mixture models are generally identifiable and can be estimated by maximum likelihood with standard asymptotics, they are known to be unreliable in the small-to-moderate sample sizes typically observed in clinical trials [16]. Moreover estimation is particularly challenging when follow-up is short and empirical survival

☆ Funding: USPHS grants U10CA21661, U10CA180822, U10CA180868; Pennsylvania Department of Health grant 4100057652.
* Corresponding author at: Department of Clinical Sciences, UT Southwestern Medical Center, Dallas, TX 75390-9066, USA.
E-mail address: dheitjan@smu.edu (D.F. Heitjan).

http://dx.doi.org/10.1016/j.cct.2017.05.012
Received 18 January 2017; Received in revised form 4 May 2017; Accepted 19 May 2017
Available online 22 May 2017

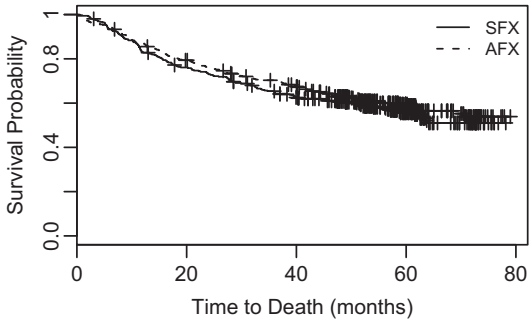

Fig. 1. Kaplan-Meier curves in RTOG 0129: SFX = standard-fractionation radiotherapy alone; AFX = accelerated-fractionation radiotherapy plus concurrent cisplatin-based chemotherapy.

curves have not had time to reach a plateau. Because predictions are of most value early in the trial, which is precisely when no subjects have extended follow-up, there is concern that the mixture-modeling approach may not be helpful in typical practice.

Another motivating example is Radiation Therapy Oncology Group trial 0129 (RTOG 0129) [1], which compared accelerated-fractionation radiotherapy plus concurrent cisplatin-based chemotherapy to standard-fractionation radiotherapy alone among patients with oropharyngeal squamous-cell carcinoma. The study found no statistically significant treatment effect on survival. An interesting feature of the data is that the Kaplan-Meier (KM) curves for both arms level off (Fig. 1), suggesting that a large fraction of subjects are effectively cured of their cancer.

In this paper we further explore the Weibull cure-mixture prediction model that provided the best fit in Chen's immunotherapy trial example. We apply this model, together with a non-cure Weibull model and a nonparametric model, to create and empirically evaluate interim predictions in RTOG 0129. We study the relative performance of the methods in a simulation experiment.

2. Methods

2.1. General framework for event-time prediction

We briefly review the prediction framework described in detail in [2-4,7,8]. Assume we are conducting a two-arm randomized trial that began enrolling patients at calendar time 0. Subjects arrive according to a Poisson process with a constant rate of μ per unit of time and are randomized 1:1 between study arms. Each participant can either i) develop the event of interest, ii) remain in the trial without occurrence of the event, or iii) become lost to follow-up. At current calendar time $t_0 > 0$ we seek to make predictions about the future course of the trial. A typical objective is to predict the calendar time T^* at which the D^*th event will occur; for example, at $t_0 = 6$ months we may predict the time T^* at which event $D^* = 100$ will occur. The essence of the method is to use the accumulating trial data to estimate the accrual/survival model, which we then use to create predictions about the future course of the trial.

We previously developed a Bayesian prediction method based on a Weibull (non-cure) survival model [4], which assumes that survival time T in arm j, $j = 0,1$ follows the survival function

$$\Pr[T > t] = S(t|\alpha_j, \beta_j) = \exp[-(t/\beta_j)^{\alpha_j}], \quad t > 0, \quad \alpha_j > 0, \quad \beta_j > 0, \tag{1}$$

and that time to loss to follow-up in arm j independently follows a Weibull with arm-specific scale and shape parameters. The first step in the prediction modeling is to specify priors for the enrollment rate and the parameters of the Weibull event and loss to follow-up distributions in each arm. To create predictions at time t_0 we compute the posterior density of the parameters using the data accrued up to that time. Then we conduct the following steps many times:

1. Sample a set of parameters from the posterior using importance sampling (or some other method).
2. Conditional on the sampled parameters, sample a data set from the predictive distribution of the enrollment, survival and loss times:
   (a) Simulate the event and loss to follow-up times for participants who are enrolled and on study but have not yet experienced an event.
   (b) If the total enrollment goal has not yet been reached, simulate the enrollment, event and loss times for a hypothetical set of participants who have not yet been enrolled.
   (c) Determine each subject's date of event or loss (real or simulated), and rank the event dates among subjects who either have had an event or are predicted to have an event.
   (d) Identify the date of the landmark time T^*.

Each replication of steps 1 and 2 generates a draw from the predictive distribution of the landmark time T^*. Repeating them many times, one can predict the landmark time as, for example, the median of the simulated distribution of T^*, with 95% prediction interval equal to the interval between the 2.5th and 97.5th centiles of the simulated distribution.

2.2. Prediction using the Weibull cure-mixture model

Common failure-time models assume that every subject is susceptible to the event of interest and will experience it if followed long enough. This assumption fails in studies where there is a possibility of cure. The cure-mixture model addresses this deficiency by positing that the study sample is a mixture of uncured individuals (who will experience the event of interest if not censored) and cured individuals (who will never experience it no longer how long we follow them) [15–18]. Let T denote the time to the event of interest. The Weibull cure-mixture model asserts that for a subject in arm j,

$$\Pr[T > t] = (1 - \rho_j) \times S(t|\alpha_j, \beta_j) + \rho_j, \tag{2}$$

where $S(t|\alpha_j,\beta_j)$ is the Weibull survival function from Eq. (1), and $\rho_j \in [0,1]$ is the probability of cure in arm j. When t is large, the survival function approaches the cure fraction ρ_j; when $\rho_j = 0$, it reduces to the standard Weibull survival model.

Many variations of the cure-mixture model appear in the literature [19]. A common version models the cure probability with a logistic regression and the survival with a Weibull distribution, the latter being a popular choice thanks to its flexibility and its similarity to Cox regression [20]. Our analyses will model both the cure probability and the Weibull survival parameters only as functions of the randomization arm (see Eq. (2)).

Under the Weibull cure-mixture model the overall prediction framework remains the same, except that one must also predict cure status for the censored and not-yet-enrolled subjects. We generate the cure status of each unenrolled participant in arm j by binomial sampling with the sampled cure probability $\widetilde{\rho}_j$. For a subject in arm j who was enrolled at calendar time e, and who did not experience an event by prediction time $t_0 \geq e$, the conditional cure probability for sampling is

$$\widetilde{\Pr}[\text{cured}|T > t_0 - e] = \frac{\widetilde{\rho}_j}{\widetilde{\rho}_j + (1 - \widetilde{\rho}_j) \times S(t_0 - e; \widetilde{\alpha}_j, \widetilde{\beta}_j)}, \tag{3}$$

where $S(\cdot;\widetilde{\alpha}_j, \widetilde{\beta}_j)$ is the Weibull survival function given sampled parameters $\widetilde{\alpha}_j$ and $\widetilde{\beta}_j$. The cure status for these already enrolled subjects in arm j is then simulated from the binomial distribution with this estimated cure probability. For a subject who is simulated to be cured, the event time is imputed as infinity. If a subject is simulated as not cured, we impute the event time by drawing from the unconditional Weibull for a new subject, or the Weibull conditional on the event time being at least $t_0 - e$ for an existing censored subject.