# CURE MODELS IN ONCOLOGY CLINICAL TRIALS

Legrand, C. and A. Bertrand

REPRINT | 2019 / 43

UCLouvain

ISBA | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

# *Krantz Template*

# Part I

# Appropriate part for Chapter on Cure Models

# 1

# *Cure models in cancer clinical trials - C. Legrand and A. Bertrand*

**CONTENTS**




*Abstract: Due to the advances in medical research in the past decades, cancer is not necessarily a fatal disease anymore. For specific cancer types, one can now reasonably expect a fraction of long-term survivors to show-up in cancer clinical trials. The presence of short-and long-term survivors may lead to a violation of the proportional hazards assumption and therefore jeopardize the use of of the popular Cox model. Furthermore, in such a situation the proportion of "cured" patients becomes a crucial component of the assessment of patients benefit, and being able to distinguish a curative from a life-prolonging effect conveys important additional information in the evaluation of a new treatment. To address these issues, specific "cure models" have been proposed in the statistical literature. In this chapter we introduce the two main fami-*

*lies of such models: mixture cure models and promotion time cure models. We elaborate on how and when to use them, and discuss that in practice, it is not only a matter of whether or not there are cured patients in the data, but that, as in classical survival analysis, the appropriate model to be used should be carefully chosen, based on the main features of the data and with a strong emphasis on the PH assumption. C. Legrand and A. Bertrand* [1]

## 1.1 Introduction

Over the last decades, it has become more and more common in cancer clinical trials to observe patients experiencing long-term relapse-free survival, and cure has become a reality for both patients and clinicians [32]. It is indeed nowadays widely accepted that for a number of cancer types, such as early-stage breast cancer [43], colon cancer [44], or childhood acute lymphocytic leukaemia [5] among others, treatment can lead to cure for a fraction of the patients.

Therefore, for these types of cancer, the primary goal of cancer therapy can not only be prolongation of survival but should shift towards cure. [32] pointed out that this is especially true for cancers occurring in children, as in this case a curative treatment will yield many years of healthy life, while a life-prolonging treatment will only offer a limited benefit before relapse takes the child's life. However, it may also be crucial for adult patients to express the benefit of new therapies not only in terms of delaying death but also in terms of cure, as this can free the patients from cancer-associated sufferings, which could sometimes be more unbearable to patients than death itself [32]. In that new paradigm, the proportion of cured patients (often referred to as the *cure rate*) is an important measure of long-term survival benefit.

A common feature of time-to-event data in clinical trials is *right censoring*, meaning that, at the time of analyzing the trial results, some of the patients have not (yet) experienced the event of interest. Right censoring unfortunately prevents us from distinguishing the cured patients from the patients who will experience the event of interest after the censoring time. Furthermore, a very common assumption is that censored subjects will follow the same survival pattern after withdrawal as non-censored subjects. This leads to the implicit assumption that, if the follow-up was long enough, one would observe the event of interest for all patients, which is obviously not the case in clinical trials of curable diseases.

The most commonly used statistical methods for the analysis of clinical trials with a time-to-event type of endpoint are most certainly the (non-parametric) logrank test [34] and the (semi-parametric) Cox proportional hazards (PH) model [16]. Although the latter is widely used, it is well known that

[1] Institute of Statistics, Biostatistics and Actuarial Sciences, LIDAM, UCLouvain

it relies on the proportional hazards assumption, i.e. while the absolute underlying hazard may vary over time, the hazard ratio (HR) between the two treatment groups remains constant. Although this PH assumption is not necessary for the computation of the logrank statistic, it is well known that it is required to achieve maximal power of the test and that it is also central to the interpretation of the results [36]. Indeed, if one can assume that the HR between the two treatment groups remains constant, then this single constant can adequately be used to summarize the difference between the two survival curves over time. If the PH assumption is not met, using the semi-parametric Cox PH model, and to some extent the logrank test, may lead to both misleading and uninterpretable conclusions, and in particular if the censoring rate is high [36]. Therefore, while these "classical" survival methods are usually appropriate in clinical settings where we expect few patients to be cured and where the primary goal is to identify treatment allowing to prolong the duration of remission [46], they have been challenged over the last years by the apparition of new treatments having different mechanisms of action on the occurrence of recurrences. Indeed, testing of therapies associated with a *delayed treatment effect* or a *rebound effect* will obviously lead to non-proportional hazards situations. But this is also the case of new therapies having a *curative effect* affecting the proportion of *cured* patients, with or without affecting the timing of occurrence of recurrences for the other patients. In this latter setting, the study sample will consist of a mixture of "*cured*" and "*uncured*" (also often called "*susceptible*") patients, the latter experiencing disease recurrence after some time following inclusion in the trial. Note that if the primary endpoint is overall survival (OS) or progression-free survival (PFS), as is often the case in oncology trials, it is clear that no patient can be cured from death; one will then be speaking of long-term survivors and it is convenient to think of these long-term survivors as (statistically) cured [35, 59, 26]. Such an heterogeneous population of short and long-terms survivors may lead to the PH assumption of the Cox model to be violated. While extensions of the conventional Cox PH model have been proposed to deal with the issue of non proportionality (e.g. inclusion of time-varying covariate effects [49])), these methods do not adequately allow one to distinguish between the curative and life-prolonging effects of a new treatment [36, 59].

Nearly 70 years ago, Boag was the first statistician to draw attention to the presence of cure in the analysis of cancer-related survival data by proposing in a seminal paper a parametric cure model allowing to estimate the cure rate as one of its parameters, and assuming a log-normal model for the failure time among the uncured patients [6]. Few years later, Berkson and Gage proposed a similar model, but considering an exponential model for the uncured patients [4]. These models were then popularized by Farewell [17, 18] who proposed in 1977 to model the cure rate as a dependent variable in a logistic regression. Since then, cure models have been a popular component of the statistical literature. While the advances of statistical research on *cure models* are closely linked to the progress made in the treatment of cancer [35], these models have

also been studied in the context of other medical applications, as well as, more broadly, in other fields (such as psychology, sociology or demography). In all these contexts, interest lies in the impact of one (or more) factor not only on delaying, but also on eradicating the event of interest for a non-negligible part of the population.

Besides medical evidence, a straightforward way to identify whether a particular dataset includes a subset of cured or long-term survivor patients is to inspect the Kaplan-Meier estimate of the survival curve. If a long and stable plateau with heavy censoring is observed in the tail, this can be considered as empirical evidence of a cure fraction [59, 47]. Cure models can then be considered as a useful alternative to the standard Cox PH models to explicitly describe the heterogeneity within the patient population, and particularly if the PH assumption is violated.

Two main families of cure regression models have been proposed: *mixture cure models* and *promotion time cure models*. Mixture cure models follow the lines of the Boag model [6]. They explicitly model the survival function of the population as a mixture of two types of patients: those who are cured and those who are not. Typically, they are composed of two sub-models: a first one for the probability of being (un-)cured, typically modeled via a logistic regression, and a second one which is a survival model for the patients who are not cured, commonly a parametric Weibull or a semi-parametric Cox PH model. As we will briefly review in Section 1.2, many variations of the mixture cure models have been proposed in the statistical literature. A major advantage of these models is the possibility to disentangle the effects of covariates, and in particular of the treatment, on the probability of cure and on the failure time of the uncured patients, resulting in a more accurate picture of the clinical benefit than with a standard Cox analysis. Promotion time cure models, also referred to as *non-mixture cure models*, are based on a totally different approach and have been originally proposed to model the biological evolution of carcinogenic cells [56, 57]. They are also sometimes called *bounded cumulative hazard models* and we will show that some specific promotion time cure models can be thought of as a Cox PH model that allows a cure fraction. A number of promotion time cure models have been proposed in the statistical literature; see Section 1.3 for a short overview. As we will discuss later, the interpretation of covariates is different with the promotion time cure models and the mixture cure models. We will also demonstrate in Section 1.4 that depending on the type of data and on the questions to be answered, a promotion time cure model or a mixture cure model may be more appropriate.

While cure models are well known in the field of statistics and have been quite extensively studied in the statistical field for at least the past 20 years, they have not reached the same popularity in a more clinical setting. Cure has now become a reality for both the patients and the clinicians in some types of cancer; however, despite the fact that cure models can therefore be an interesting way to characterize and study patients survival, they are still an a underused statistical tool in the context of oncology trials. This may be due

to the extreme popularity of the Cox PH model, to the lack of implementation of cure models in standard software, and probably also to a lack of knowledge about these models in this setting. The purpose of this chapter is therefore to introduce the main ideas regarding the use of cure models for survival data analysis in the framework of oncology clinical trials. For a larger-scale and more technical overview of these models, we refer the reader to [40] and [1] and the references therein.

In this chapter, we focus on right-censored data, unless otherwise specified. We will denote by $T$ the time to the event of interest, with $F(t)$ and $S(t) = P(T > t) = 1 - F(t)$ the associated distribution and survival functions, respectively. Let $C$ be the right-censoring time, with distribution function $G(t)$. We therefore observe $\tilde{T} = \min(T, C)$ and the censoring indicator $\delta = I(T \leq C)$. In Sections 1.2 and 1.3, we give some more details about the two main families of cure models. The motivation on when to use these models, as well as issues related to the choice of the model, are discussed and illustrated with simulations results in Section 1.4. One freely available real database is analyzed in Section 1.5, which allow us to discuss further the interpretation of cure models.

## 1.2 Mixture cure models

### 1.2.1 Model and properties

The idea behind mixture cure models is to explicitly take into account the fact that the population is a mixture of two groups of patients: the cured patients, who will never experience the event of interest, and the uncured (susceptible) ones. If we denote by $Y$ the cure indicator, with $Y = 1$ corresponding to a susceptible patient and $Y = 0$ otherwise, we can define the probability of being uncured (or susceptible) as $\pi = P(Y = 1)$. Assuming that, for cured patients, the survival function is $S_c(t) = P(T > t \mid Y = 0) = 1$ for all $t$ (i.e., a degenerate survival function), it is natural to define the mixture cure model by the following unconditional survival function of $T$:

$$S_{pop}(t) = P(T > t) = (1 - \pi) + \pi S_u(t), \tag{1.1}$$

where the subscript *pop* indicates that the survival function relates to the whole population and $S_u(t) = P(T > t \mid Y = 1)$ is the survival function of the susceptible patients, which is proper (i.e., $\lim_{t\to\infty} S_u(t) = 0$).

The mixture cure model therefore appears as a combination of two sub-models, one for the probability of cure (often referred to as the *incidence model*) and one for the survival of the uncured patients (the *latency model*). Each of these sub-models can be allowed to depend on (potentially different)

covariates, and the mixture cure model hence allows one to disentangle the effect of covariates on the incidence and on the latency. Given a first set of covariates $\boldsymbol{X}$, and a second set of covariates $\boldsymbol{Z}$, which might be identical to $\boldsymbol{X}$, or partially or completely different from $\boldsymbol{X}$, the mixture cure model (1) then writes the population survival function as

$$S_{pop}(t \mid \boldsymbol{X}, \boldsymbol{Z}) = (1 - \pi(\boldsymbol{X})) + \pi(\boldsymbol{X}) S_u(t \mid \boldsymbol{Z}). \tag{1.2}$$

With such a mixture cure model, the covariates, and in particular the treatment indicator, are therefore allowed to have dissimilar effects on the probability of cure and on the timing of events for the susceptible individuals. This obviously carries additional information about the type of treatment effect compared to a standard Cox PH model. Furthermore, the sets of covariates $\boldsymbol{X}$ and $\boldsymbol{Z}$ may be different, which is in line with the intuition that medical and patient related factors associated with short- or long-term effects are not necessarily the same.

Most of the time, the impact of covariates on the incidence is modeled via a logistic regression model, as originally proposed by [17]. The vector of covariates $\boldsymbol{X}$ and the corresponding vector of parameters $\boldsymbol{\gamma}$ then contain an intercept, and the logistic incidence model for the probability of being uncured can be written

$$\pi(\boldsymbol{X}) = \frac{\exp(\boldsymbol{X}^T \boldsymbol{\gamma})}{1 + \exp(\boldsymbol{X}^T \boldsymbol{\gamma})}. \tag{1.3}$$

Few alternatives to the logistic regression have been proposed and without real implementation in practice. Existing ideas consist in considering a more flexible modeling of the incidence, for example using splines [54] or a single-index structure [2].

On the other hand, various ways to model the latency have been considered and applied in the literature. We may distinguish the parametric and the non- (or semi-) parametric mixture cure models. In the former, the survival times of uncured patients follow a parametric model, while the latter leaves the baseline survival function of the uncured patients unspecified. Fully non-parametric latency models have also been proposed but are not in use in the medical literature (see for example [48] for the case of no covariate and [37] for the case with covariates).

Several parametric models have been considered for the latency (a review can be found in [1]), the most popular being probably the Weibull [18] PH model with

$$S_u(t \mid \boldsymbol{Z}) = \exp\left(-\lambda \exp(\boldsymbol{Z}^T \boldsymbol{\beta}) t^{\rho}\right), \tag{1.4}$$

with $\boldsymbol{\beta}$ the vector of parameters corresponding to $\boldsymbol{Z}$, $\lambda > 0$ the shape parameter and $\rho > 0$ the scale parameter ($\rho = 1$ in the case of an exponential model, [19]). The hazard function corresponding to this survival function is then

$$\lambda_u(t \mid \boldsymbol{Z}) = \lambda \rho \exp(\boldsymbol{Z}^T \beta) t^{\rho - 1}. \tag{1.5}$$

The Weibull model is fairly flexible and is often considered to provide a good description of survival times in biomedical applications. However, the assumption of a monotone baseline hazard function may also be problematic. The main advantage of these fully parametric cure models, as we will see in Section 1.2.4, lays in the simplicity of the estimation procedure. Also, it may be useful to rely on parametric models if we are interested in modeling the survival time in the investigated treatment arms, and not only the relative difference between two survival curves [36].

Such parametric models may not be robust to the violation of the distributional assumption for the survival times of the uncured patients. If we still want to rely on a parametric estimation procedure, a possibility is to consider a very flexible parametric form for the baseline hazard function, such as a piecewise constant baseline hazard or the use of splines [14]. However, semi-parametric alternatives, in which the baseline hazard of the latency model is left totally unspecified, have gained popularity in the statistical literature, despite the fact that obtaining semi-parametric estimators and their standard errors can be computationally challenging.

A common alternative to parametric models is indeed to consider a semi-parametric Cox PH model for the latency. This model has been introduced by [25] and has then been extensively studied in the literature [48, 38, 47, 30, 14]. Such an approach has the advantage of leaving the baseline hazard function $\lambda_{uo}(\cdot)$ of the uncured patients unspecified,

$$\lambda_u(t \mid \boldsymbol{Z}) = \lambda_{uo}(t) \exp(\boldsymbol{Z}^T \boldsymbol{\beta}),$$

with the corresponding survival function

$$S_u(t \mid \boldsymbol{Z}) = S_{u0}(t)^{\exp(\boldsymbol{Z}^T \boldsymbol{\beta})}. \tag{1.6}$$

A very important feature of the (semi-)parametric logistic/PH mixture cure model is that the PH assumption is made at the level of the uncured patients, but not at the level of the population. This appears clearly on Figure 1.1 which displays examples of population survival functions from parametric logistic/Weibull PH mixture cure models including a binary variable (representing for example the treatment group). On each plot, the solid line represents the recurrence free survival (RFS) in the control group, with a cure rate of 40%. The dotted line represents the RFS achieved with an experimental treatment (a) which only has a long-term effect, by increasing the proportion of cure (incidence) but with no impact on delaying recurrences amongst the uncured patients (latency); (b) which only has a short-term effect, by delaying recurrences amongst the uncured patients but with no impact on the proportion of cure; (c) which has both a short- and a long-term effect, affecting both the proportion of cure and the time of recurrences for the uncured patients. As we can see, a curative treatment acting only on the proportion of cure still leads to a PH situation for the population, with the two curves attaining their plateaus by "running parallel to each other" [46]. On the other hand,

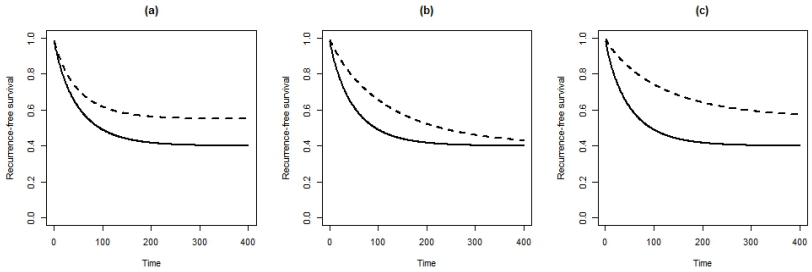



**FIGURE 1.1**
Survival functions from logistic/Weibull PH models for two treatment groups. The control treatment (solid line) leads to 40% cure; while the experimental treatment (dotted line) has (a) a long-term effect, i.e. an impact on the probability of cure (incidence); (b) a short-term effect, i.e., an impact on the latency; (c) both a short- and long-term effect.

the other two situations clearly lead to a violation of the PH assumption for the population.

As in classical survival analysis, the Accelerated Failure Time (AFT) and the PH models often confront each other in cure analysis. In the logistic/AFT mixture cure model, the survival times $T^u$ of the uncured patients are modeled as

$$\log(T^u) = \boldsymbol{Z}^T\boldsymbol{\beta} + \sigma\epsilon, \tag{1.7}$$

with $\sigma > 0$ a scale parameter and $\epsilon$ an error term whose density function $f_\epsilon$ can be specified in a parametric way or left unspecified, leading to respectively parametric [58, 39, 45] or semi-parametric [28, 62] logistic/AFT models. These models do not make any PH assumption, neither at the level of the uncured patients (except when a Weibull distribution is assumed for $\epsilon$) nor at the population level. They are however rarely used in the medical literature.

### 1.2.2 Interpretation

A major advantage of mixture cure models is the fact that each set of parameters can be interpreted separately. As mentioned earlier, the effect of the covariates on both components of the model (incidence and latency) can therefore be clearly disentangled. By quantifying separately the effect of the treatment on the probability of being a long-term survivor and on the event times for those who are not, this model allows one to distinguish a curative from a life-prolonging treatment effect.

Furthermore, the commonly used logistic/(semi-)parametric PH model leads to odds and hazard ratios (OR and HR) whose interpretation is well

known. In the logistic regression incidence sub-model, the parameters $\boldsymbol{\gamma}$, representing the impact of the covariates $\boldsymbol{X}$ on the probability to be uncured $\pi(\boldsymbol{X})$, are interpreted as usual. Parameter values above 0 are associated with covariates which increase the risk to be uncured (and therefore decrease the risk to be cured) when their value increases, with the reverse for parameter values below 0. The quantity $1-\pi(\boldsymbol{X})$ can be interpreted as the cure rate for patients with covariate value $\boldsymbol{X}$.

The parameters $\boldsymbol{\beta}$ associated with the covariates $\boldsymbol{Z}$ in the latency model represent the impact of the covariates on the time to event for uncured patients and are interpreted according to the model used. With a (semi-)parametric PH model, a positive value is associated to a covariate which increases the hazard of events (and therefore accelerates the events) when its value increases; with the reverse for values below 0. In an AFT model, the covariates act multiplicatively on the time; a positive coefficient indicates a longer time-to-event when the value of the covariate increases.

### 1.2.3 Identifiability

A general and informal rule that holds for all cure models requires the follow-up of the study to be sufficiently long: the estimated survival function should exhibit a long plateau containing many censored observations [1]. More formally, the maximum possible event time should be smaller than the maximum possible censoring time.

In a semi-parametric mixture cure model, the latency component $S_u(\cdot)$ (or part of it) is left unspecified and estimated non-parametrically. In this case, some additional information is required for identifiability. In [47] and [48], the survival function (or the baseline survival function) is constrained to reach 0 at the largest observed event time. This *zero-tail constraint* seems natural in contexts in which a cure model is justified, i.e., when cured subjects are known to exist and the follow-up is sufficiently long after the largest event time [48].

### 1.2.4 Model estimation

The estimation of mixture cure models is classically based on the maximization of the likelihood function. Assume we have i.i.d. data $(\tilde{T}_i, \delta_i, \boldsymbol{X}_i, \boldsymbol{Z}_i), i = 1, ..., n$, where $n$ is the number of patients, with $\tilde{T}_i = \min(T_i, C_i)$ the observed time and $\delta_i = I(T_i \leq C_i)$ the censoring indicator. We have to keep in mind that the cure status is only observed for uncensored observations (who are obviously uncured), not for the censored ones. As in standard survival analysis, the likelihood function is based on the contributions of two types of observations: the uncensored ones ($\delta = 1$), all corresponding to uncured patients (occurring with probability $\pi(\boldsymbol{X})$), and the censored ones ($\delta = 1$), corresponding either to cured patients (with probability $1-\pi(\boldsymbol{X})$) or to uncured patients

(with probability $\pi(\boldsymbol{X})$). The likelihood function is then defined as

$$\begin{aligned} L_{MCM} = \quad & \prod_{i=1}^{n} \quad \left[\pi(\boldsymbol{X}_i) f_u(\tilde{T}_i \mid \boldsymbol{Z})\right]^{\delta_i} \\ & \times \prod_{i=1}^{n} \quad \left[1 - \pi(\boldsymbol{X}_i) + \pi(\boldsymbol{X}_i) S_u(\tilde{T}_i \mid \boldsymbol{Z})\right]^{1-\delta_i} \end{aligned} \tag{1.8}$$

with $f_u(\tilde{T}_i \mid \boldsymbol{Z}) = -(d/dt) S_u(\tilde{T}_i \mid \boldsymbol{Z})$.

For fully parametric mixture cure models, the parameters can be estimated by maximizing the likelihood function (1.8) via numerical techniques, such as the Newton-Raphson algorithm. Asymptotic standard errors can be obtained by inverting the Fisher information matrix of second order derivatives of $\log(L_{MCM})$. Some adaptations of this likelihood maximization have been proposed for more flexible parametric models; see for example [58] who propose a two-step maximization in the case of a logistic/AFT model with an error distribution from the extended family of generalized gamma.

As mentioned earlier, the semi-parametric logistic/Cox PH mixture cure model does not satisfy the PH assumption at the population level and the partial likelihood approach developed by [16] cannot be applied to estimate this model. Indeed, since the survival function of the uncured patients is conditional on the cure status, one can not eliminate $S_0(t|Y = 1)$ in the Cox PH mixture cure model likelihood without losing information about $\boldsymbol{\gamma}$ [15]. Several estimation procedures have been presented in the literature. Given that the cure status is a latent variable, [38] and [47] have proposed to rely on the Expectation-Maximization (EM) algorithm. An interesting feature of this approach is that the complete-data likelihood, obtained from the explicit contributions of the uncensored observations ($\delta = 1, Y = 1$), censored and uncured observations ($\delta = 0, Y = 1$) and censored and cured observations ($\delta = 0, Y = 0$), can be factorized into two elements. Each element depends only on the parameters of one of the two parts of the model. This obviously simplifies the maximization in the M-step, as it can be performed separately for each set of parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$. [15] advises to use non-parametric bootstrap to obtain the standard errors of the estimated parameters. Other proposed estimation approaches include methods based on a marginal likelihood, obtained by integrating out the likelihood function (1.8) over the distinct event times ($\tilde{Y}_{(j)}, j = 1, ..., r$) using Monte-Carlo approximation [25]. A penalized likelihood approach, approximating the baseline conditional hazard by a linear combination of cubic normalized B-splines, has been introduced by [14].

The estimation procedures proposed until now for the semi-parametric logistic/AFT model are all based on the EM algorithm and follow the same ideas as for the semi-parametric logistic/Cox PH mixture cure model, except for the latency estimation in the M-step, which is then based on extensions of the methods proposed for the classical semi-parametric AFT models.

We refer to [1] for a more detailed review of these estimation methods as well as a discussion on the estimation methods for non-parametric mixture cure models.

### 1.2.5 Model implementation

The main `R` package for the estimation of mixture cure models is the `smcure` package developed by [11] which allows one to fit a wide range of semi-parametric mixture cure models. Available models for the incidence include the logistic regression model but also other generalized linear models with various link functions, such as the probit. Both semi-parametric PH and AFT models can be fitted for the latency part. The estimation procedure is based on the EM algorithm as developed by [38] and [47] for the PH case and by [62] for the AFT case. The variance of the estimated parameters is obtained via bootstrap.

A freely available `SAS` macro, called `PSPMCM`, has been developed for the parametric and semi-parametric logistic/PH mixture cure model by [15]. The maximization of the likelihood function is performed using the Newton-Raphson procedure (as implemented in `PROC NLMIXED`) when the latency is modeled via a parametric model (exponential, Weibull, lognormal or log-logistic) and through an EM algorithm for the semi-parametric case. Standard errors of the estimators are obtained either by inverting the observed Fisher information matrix at the last iteration or via non-parametric bootstrap.

## 1.3 Promotion time cure models

### 1.3.1 Model and properties

In promotion time cure models, the existence of a cure fraction is taken into account by directly choosing an improper form for the survival function of the whole population, instead of separately modeling the survival of cured and uncured patients. This corresponds to considering that cured patients have an infinite survival time. Given a vector of covariates $\boldsymbol{X}$, these models, introduced by [56] and [57], have the form

$$S_{pop}(t|\boldsymbol{X}) = \exp\left\{-\theta(\boldsymbol{X})F(t)\right\}, \tag{1.9}$$

where $F(\cdot)$ is the (proper) cumulative distribution function (cdf) of some nonnegative random variable such that $F(0) = 0$ and $\theta(\boldsymbol{X})$ is a known link function with an intercept. The baseline cdf $F(\cdot)$ can either be modeled parametrically, yielding *parametric* promotion time cure models (as is done, among others, by [12]), or left unspecified, which leads to *semi-parametric* promotion time cure models (see, for example, [51]). The cumulative hazard function of this model is $\theta(\boldsymbol{X})F(t)$, which is bounded: for this reason, promotion time cure models are also sometimes called *bounded cumulative hazard models*.

In model (1.9), the cure probability is

$$\lim_{t\to\infty} S_{pop}(t|\boldsymbol{X}) = \exp\left\{-\theta(\boldsymbol{X})\right\}. \tag{1.10}$$

The hazard function corresponding to the survival function (1.9) is

$$h_{pop}(t|\boldsymbol{X}) = \theta(\boldsymbol{X})f(t),$$

where $f(t) = dF(t)/dt$ is the baseline density function. Contrary to the mixture cure model considered in the previous section, the promotion time cure model possesses the proportional hazards property:

$$\frac{h_{pop}(t|\boldsymbol{X}_i)}{h_{pop}(t|\boldsymbol{X}_j)} = \frac{\theta(\boldsymbol{X}_i)}{\theta(\boldsymbol{X}_j)}$$

and is therefore sometimes referred to as the *proportional hazards cure model*.

An example of survival functions from a promotion time cure model are displayed and discussed later in Section 1.3.3.

### 1.3.2 Link with other models

The semi-parametric promotion time cure model with an exponential link function can actually be seen as a generalization of the Cox PH model [42]. If we assume the link function $\theta(\boldsymbol{X}) = \exp(\beta_0 + \boldsymbol{X}^T\boldsymbol{\beta})$ with intercept $\beta_0$, then Equation (1.9) becomes

$$\begin{aligned} S_{pop}(t|\boldsymbol{X}) &= \exp\Big\{-\exp(\beta_0 + \boldsymbol{X}^T\boldsymbol{\beta})F(t)\Big\} \\ &= \exp\Big\{-\exp(\boldsymbol{X}^T\boldsymbol{\beta})\exp(\beta_0)F(t)\Big\} \\ &= \exp\Big\{-\exp(\boldsymbol{X}^T\boldsymbol{\beta})H(t)\Big\}, \end{aligned}$$

i.e., a PH model in which $H(t) = \exp(\beta_0)F(t)$ is a *bounded* cumulative hazard function, taking values in $[0, \exp(\beta_0)]$. In the Cox PH model, the cumulative hazard function is not bounded [16]. However, since, in practice, the estimator of the cumulative hazard function of a Cox PH model is bounded, we have the following links between the estimates obtained from a promotion time cure model with exponential link, $\widehat{\beta}_{0,PT}$, $\widehat{\boldsymbol{\beta}}_{PT}$ and $\widehat{F}_{PT}(t)$, and the estimates from a Cox PH model estimated by maximizing the profile likelihood, $\widehat{\boldsymbol{\beta}}_{PH}$ and $\widehat{H}_{PH}(t)$:

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_{PT} &= \widehat{\boldsymbol{\beta}}_{PH} \\ \exp(\widehat{\beta}_{0,PT}) &= \widehat{H}_{PH}(T_{(n)}) \\ \exp(\widehat{\beta}_{0,PT})\widehat{F}_{PT}(t) &= \widehat{H}_{PH}(t), \end{aligned}$$

where $T_{(n)}$ is the largest event time.

This means that, when the exponential link function is used, the estimates of the regression coefficients of a semi-parametric promotion time cure model can, in practice, be obtained from fitting a standard Cox PH model. This

explains that, as long as the PH assumption is met, the Cox model also provides reliable results even in the presence of a non-negligible cure fraction. In that situation, however, the parameters should be interpreted in the context of a promotion time cure model, i.e., taking into account that cure is a possibility.

Although model (1.9) is not equivalent to a mixture cure model, it is possible to re-express the population survival function with a mixture expression as

$$\{1 - p(\boldsymbol{X})\} + p(\boldsymbol{X})S_u(t|\boldsymbol{X}) \tag{1.11}$$

where $p(\boldsymbol{X})$ is the probability of being susceptible, and $S_u(\cdot|\boldsymbol{X})$ is the conditional survival function of the uncured patients. We have already seen that in the promotion time model, the cure probability is given by

$$1 - p(\boldsymbol{X}) = \exp\{-\theta(\boldsymbol{X})\}, \tag{1.12}$$

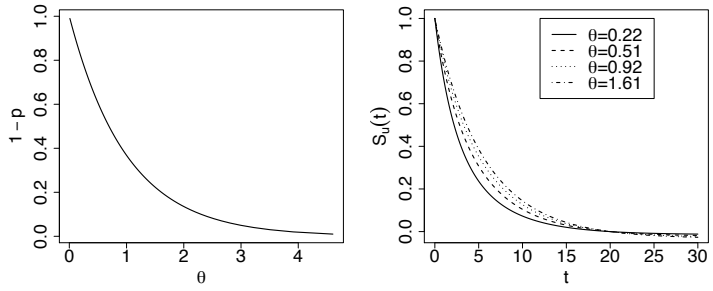
and one can show that the conditional survival of the susceptible subjects is [12]

$$S_u(t|\boldsymbol{X}) = \frac{\exp\{-\theta(\boldsymbol{X})F(t)\} - \exp\{-\theta(\boldsymbol{X})\}}{1 - \exp\{-\theta(\boldsymbol{X})\}}. \tag{1.13}$$

However, we clearly see from equations (1.12) and (1.13) that, in this formulation of the promotion time cure model as a mixture cure model, both the incidence and the latency depend on the same set of covariates $\boldsymbol{X}$, whereas this is not necessarily the case in a mixture cure model. Moreover, these covariates appear at more than one place in the conditional survival function (1.13) which is never the case in the classical versions of the mixture cure model.

### 1.3.3 Interpretation

As mentioned previously, in the promotion time cure model, the covariates $\boldsymbol{X}$ affect both the probability of being cured and the survival of uncured patients. This is best understood by considering the seminal biological interpretation of this model, which was developed with the idea of modeling cancer relapse [12]. Assume that, after an initial treatment, individual $i$ has $N_i$ carcinogenic cells that are left active, i.e., cells which could metastasize. $N_i$ is then an unobservable latent variable assumed to follow a $Poisson(\theta(\boldsymbol{X}_i))$ distribution. The cured individuals are those for whom $N_i = 0$. The $k$th carcinogenic cell ($k = 1, \ldots, N_i$) takes a time $W_{ik}$ (called the *promotion time*) to produce a detectable tumor mass. Conditionally on $N_i$, the variables $W_{ik}$, $k = 1, \ldots, N_i$ are mutually independent, independent of $N_i$ and have a common cdf $F(t)$. The time until the relapse of the cancer, which is the observed event time, is




(a) (b)

**FIGURE 1.2**
Representation of the effect of the value of $\theta$ on the cure probability and on the conditional survival function. (a) Cure probability as a function of the value of $\theta$; (b) Conditional survival function for different values of $\theta$, corresponding to cure probabilities of 0.20 (when $\theta = 1.61$), 0.40 (when $\theta = 0.92$), 0.60 (when $\theta = 0.51$) and 0.80 (when $\theta = 0.22$).

then $T_i = \min(W_1, \ldots, W_{N_i})$, which has survival function:

$$\begin{aligned} S(t|\boldsymbol{X}) &= P(N = 0|\boldsymbol{X}) + \sum_{j=1}^{\infty} P(T_1 > t, \ldots, T_j > t|N = j, \boldsymbol{X}) P(N = j|\boldsymbol{X}) \\ &= \exp\{-\theta(\boldsymbol{X})\} + \sum_{j=1}^{\infty} \{1 - F(t)\}^j \frac{\theta(\boldsymbol{X})^j \exp\{-\theta(\boldsymbol{X})\}}{j!} \\ &= \exp\{-\theta(\boldsymbol{X})\} + \exp\left[\{1 - F(t)\}\theta(\boldsymbol{X}) - \theta(\boldsymbol{X})\right] - \exp\{-\theta(\boldsymbol{X})\} \\ &= \exp\{-\theta(\boldsymbol{X})F(t)\}, \end{aligned}$$

which corresponds to Equation (1.9). The covariates $\boldsymbol{X}$ have an impact on the number of cells which can metastasize: as a consequence, these covariates directly influence the cure probability, but also the conditional survival of the uncured patients.

The relation between $\theta(\boldsymbol{X})$ and the cure probability is illustrated in Figure 1.2(a), while Figure 1.2(b) displays the relation between $\theta(\boldsymbol{X})$ and the conditional survival of the uncured patients. The resulting effect on the population survival function is represented in Figure 1.3.

The parameters of a promotion time cure model can hence be interpreted based on the biological interpretation of the model. A covariate whose increase yields an increase in $\theta(\boldsymbol{X})$ actually increases the mean number of cells which can metastasize; larger values of this covariate are hence associated with a lower cure probability and a lower survival (at all times) for the uncured patients. Conversely, if an increase in a covariate lowers the value of $\theta(\boldsymbol{X})$,

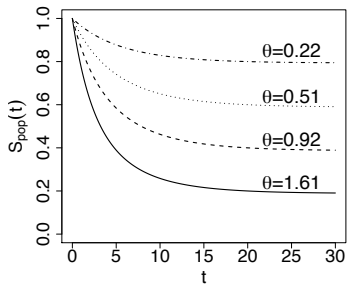


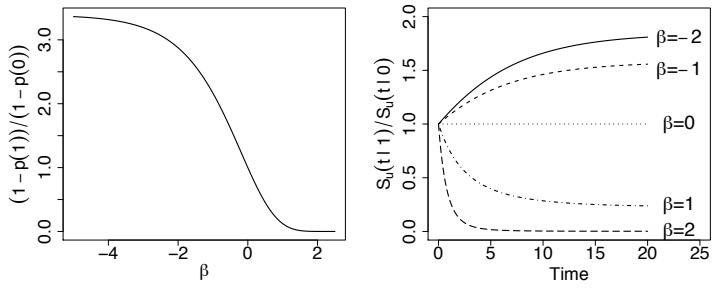

**FIGURE 1.3**
Survival function for different values of $\theta$, corresponding to cure probabilities of 0.20 (when $\theta = 1.61$), 0.40 (when $\theta = 0.92$), 0.60 (when $\theta = 0.51$) and 0.80 (when $\theta = 0.22$).

then larger values of this covariate are associated with a higher cure probability and a better survival (at all times) for the uncured patients.

In the (common) particular case where $\theta(\boldsymbol{X})$ is assumed to be $\exp(\beta_0 + \boldsymbol{X}^T\boldsymbol{\beta})$, then a one-unit increase in a continuous covariate $X_1$ is associated with a multiplication of the mean number of cells that can metastasize by a factor $\exp(\beta_1)$. For a binary covariate $X_2$ (1 = treatment versus 0 = control), the mean of $N$ is $\exp(\beta_2)$ times larger in the treatment group than in the control group.

The biological interpretation thus allows one to easily understand the different components appearing in Equation (1.9) and the covariate effect. However, promotion time cure models can also be used in contexts where such a biological interpretation does not hold. For example, [7] explain that, in a fertility progression study where the transition to second or third birth is analyzed, arguments in favor of the conception of another child can be seen as the latent cells. The time needed for these arguments to be convincing is then the observed time.

Although Equation (1.13) makes it clear that the model parameters are not easily interpreted in terms of the covariate effect on the conditional survival of the susceptible patients, such an interpretation can always (although not directly) be recovered. One example is given in Figure 1.4, in the case of a model with a binary covariate. Figure 1.4(a) contains the representation of the value of the ratio of the cure probabilities, $1 - p(1)$ and $1 - p(0)$, as a function of the coefficient of the covariate. As can be expected from our previous discussion, a negative coefficient for $X$ corresponds to a higher cure rate in the treatment group than in the control group, while the situation is reversed when the coefficient is positive. The curve is decreasing: the larger the coefficient, the smaller the cure probability in the treatment group (compared to the control group). The interpretation of the treatment effect on the conditional survival function of the susceptible patients can be found in Fig-




(a) (b)

**FIGURE 1.4**
Representation of the effect of a binary covariate in the promotion time cure model $S_{pop}(t|X) = \exp(-\exp(0.2+\beta X)F(t))$ with $F(\cdot)$ the cdf of a truncated exponential distribution of mean 6, truncated at 20. (a) Effect of the binary covariate on the ratio of cure probabilities, as a function of $\beta$. (b) Effect of the binary covariate on the ratio of survival functions, for different values of $\beta$.

ure 1.4(b). Here again, a negative coefficient for $X$ implies a more favorable situation for the treatment group, compared to the control group: the survival function for the uncured patients of the treatment group is higher, at all times, than the curve of the uncured patients of the control group.

### 1.3.4 Identifiability

The general principles regarding identifiability in cure models discussed in Section 1.2.3 still hold. In the semi-parametric promotion time cure model, the cdf $F(\cdot)$ is left unspecified and estimated non-parametrically. As was the case in the mixture cure model, identifiability requires additional information. The conditions required for the semi-parametric version of Model (1.9) with $\theta(\boldsymbol{X}) = \eta(\beta_0 + \boldsymbol{X}^T\boldsymbol{\beta})$ to be identifiable in the regression parameters and in $F$ can be found in [41], who focus on right-censored data. In addition to classical conditions about the covariates and the function $\eta(\cdot)$, [41] explain that we need a threshold $\tau$, called the *cure threshold*, such that all censored individuals with a censoring time larger than this threshold are treated as known to be cured, i.e., $T_i = C_i = \tilde{T}_i = \infty$. However, the estimation methods for this model mentioned in Section 1.3.5 force the estimated baseline cdf to be 1 beyond the largest observed failure time. This amounts to considering that no event can occur after this time: the cure threshold is then determined to be this largest event time. This restriction is the equivalent of the zero-tail constraint imposed in the mixture cure model, explained in Section 1.2.3.

### 1.3.5 Model estimation

Promotion time cure models have been studied mainly (but not only) in the Bayesian literature, because the posterior distribution of the regression parameters is often proper when using non-informative improper priors, an attractive property in a Bayesian setting [60]. Most of the Bayesian methods proposed rely on Markov Chain Monte Carlo (MCMC) methods to approximate the posterior distribution of the parameter of interest, see for example [12] in a parametric case and [23, 53] in a semi-parametric setting. A Bayesian approach including a smoothing parameter to control the degree of parametricity in the right tail of the baseline survival distribution $F(t)$ can be found in [22].

Several frequentist estimation procedures have been proposed for the semi-parametric version of Model (1.10) applied to right-censored data, among which the maximization of a profile likelihood [51, 41] and the maximization of the full likelihood through a profiling approach [61] or a backfitting approach [31].

The full log-likelihood of Model (1.9) is

$$\begin{aligned}\ell_{PTM} = &\sum_{i=1}^{n} \delta_i I\left\{-F(\tilde{T}_i)\theta(\mathbf{X}_i) + \log F\{\tilde{T}_i\} + \log\theta(\mathbf{X}_i)\right\} \\ &+ (1-\delta_i) I(\tilde{T}_i < \infty)\left\{-F(\tilde{T}_i)\theta(\mathbf{X}_i)\right\} - I(\tilde{T}_i = \infty)\theta(\mathbf{X}_i),\end{aligned}$$

where $F\{\tilde{T}_i\}$ is the jump size of the step function $F$ at $\tilde{T}_i$.

Note that, in the specific case where $\theta(\boldsymbol{X}) = \exp(\beta_0 + \boldsymbol{X}^T\boldsymbol{\beta})$ is used, the parameters can be estimated by maximizing the profile likelihood, as in a classical Cox PH model (see Section 1.3.2).

### 1.3.6 Model implementation

Two `R` packages (available on the CRAN website) contain estimation procedures for the semi-parametric promotion time cure model: `miCoPTCM`, for right-censored data (implementing the method of [31]), and `intercure`, for interval-censored data (implementing the profile likelihood approach of [29]). The `coxph` function of package `survival`, implementing the maximization of the profile likelihood, can be used to estimate the version of the promotion time cure model which is equivalent to a classical PH model.

In Stata, the `cureregr` command fits parametric cure models, the `stpm2` [3] command enables the estimation of flexible parametric cure models, and the `strsnmix` [26] command fits parametric non-mixture cure models.

### 1.3.7 Extensions

A common and important extension of Model (1.9) consists in making the baseline cdf depend on covariates:

$$S_{pop}(t|\boldsymbol{X},\boldsymbol{Z}) = \exp\left\{-\theta(\boldsymbol{X})F(t|\boldsymbol{Z})\right\}, \tag{1.14}$$

as considered, for example, by [52]. Contrary to the covariates appearing in the classical model (1.9), these new covariates $\boldsymbol{Z}$ affect the time needed by a metastatic cell to produce a detectable tumor mass, but not the mean number of such cells. As a results, these covariates influence the short-term survival, but not the cure rate. Including two sets of covariates hence allows one to distinguish between the effect on the cure rate (long-term effect), and the effect on the survival (short-term effect). However, these models do not possess the PH property, and some restrictions on the covariates may be required for identifiability (see [8] for an identifiability result in one such model).

Other possible extensions of the promotion time cure model (1.9) directly motivated through the biological interpretation of this model have also been proposed. See for example [61] who relax the assumption of mutual independence between the promotion times $Z_{ik}$ via the introduction of a subject-specific frailty term; [13] and [20], who consider other types of distribution for the number $N$ of cells that can metastasize and a number $r > 1$ of cells to be promoted to have an event; and [50] who, in addition, suggest to include covariates affecting the mean number of cells, $N$, as well as the distribution of the promotion time of each cell.

Finally, it is interesting to note that general classes of cure models have been developed, which encompass both mixture and promotion time cure models as special cases. For example, [60] propose a new class of cure models through a Box-Cox transformation of the survival function of the population, with as special cases the mixture cure models and the promotion time cure models with covariates in the baseline cdf.

## 1.4 When to use a cure model

Two main questions usually come up when speaking about cure models. First, when should we use a cure model to analyze our data or, put slightly differently, what are the consequences of not taking the cure fraction into account? Second, if we actually are in a situation where a cure model is appropriate, should we rather use a mixture cure model or a promotion time model? To discuss these questions, we present here a short simulation study investigating the consequences of misspecifying the model to be used for estimation, such as assuming a classical Cox PH model when there is actually a cure fraction in the sample, or assuming a mixture cure model when the data actually follow a promotion time cure model.

### 1.4.1 Presentation of the simulations

In this simulation study, we consider six different settings based on the way the data were generated. For each setting, 500 datasets, each containing 500

patients, were simulated. These patients are first randomly allocated to one of the treatment arms according to a binary covariate $X \sim Bern(0.5)$ ($X = 1$ for treatment versus $X = 0$ for control). Their time to event is then generated according either to a parametric PH model (no cure), a logistic/parametric PH mixture cure model or a promotion time cure model as described below. We consider right censoring with censoring times following a truncated Weibull distribution whose parameters were chosen to achieve the desired level of censoring.

**Settings 1 and 2:** The times to events are generated from a parametric PH model with an exponential baseline hazard (shape parameter set to 6). The regression parameter for the treatment indicator was set to $-1$, corresponding to a HR of 0.37 in favor of the treatment group. For these data, there are therefore no cured patients. We consider both a setting with a (relatively) high censoring rate (Setting 1) and a setting with a (relatively) low censoring rate (Setting 2).

**Setting 3:** The times to events are generated from a parametric promotion time cure model with exponential link and exponential baseline cdf (shape parameter set to 6, distribution truncated at 20). The regression parameters are set to 0.2 for the intercept and $-0.5$ for the treatment effect.

**Settings 4, 5 and 6:** The times to events are generated from a logistic/parametric PH mixture cure model with a (truncated) exponential baseline cumulative hazard function in the latency (shape parameter of 4 for the three settings and truncation limit of respectively 50, 20 and 50). These three settings differ by the inclusion of the covariate effect either in both the incidence and latency parts (Setting 4: parameter values of 1 and $-1$ for the intercept and the treatment effect in the incidence and parameter value of $-1$ for the treatment effect in latency), in the incidence part only (Setting 5: parameter values of 1 and $-1$ for the intercept and the treatment effect in the incidence and parameter value of 0 for the treatment effect in latency), or in the latency part only (Setting 6: parameter values of 0.5 and 0 for the intercept and the treatment effect in the incidence and parameter value of $-1$ for the treatment effect in latency). These settings therefore correspond to a treatment being both curative and life-prolonging (setting 4), only curative (setting 5) with no effect on the time of events amongst the uncured, or only life-prolonging (setting 6) while ultimately not impacting the long-term outcome of the patients.

The average censoring and cure rates per treatment arm and overall are presented in Table 1.1; Setting 6 is the only one leading to the same proportion of cure in both treatment arms. Figure 1.5 represents, for each setting, the theoretical survival curves by treatment group at the level of the population (obtained without considering censoring) as well as these KM estimated survival curves for one random dataset of each setting (accounting for censoring).

As expected, the PH assumption is met in settings 1, 2, and 3 but one could also consider that setting 5 meets this assumption since both curves are first parallel and then reach their plateaus at the same time. The presence of a plateau (and therefore of a cure fraction) is clear from the estimated survival curves from settings 3 to 6. Setting 2 obviously shows no plateau in the estimated survival curves; however, setting 1 may be confusing. It is therefore important to keep in mind that it is the combination of a sufficiently long follow-up, a long plateau and a sufficient number of censored observations in the plateau that can be considered as an indication of the presence of a cure fraction.

**TABLE 1.1**
Simulation settings characteristics

| | Average censoring rate | | | Average cure rate | | |
|---|---|---|---|---|---|---|
| | Overall | $X = 0$ | $X = 1$ | Overall | $X = 0$ | $X = 1$ |
| Setting 1 | 0.53 | 0.39 | 0.67 | | | |
| Setting 2 | 0.27 | 0.15 | 0.39 | | | |
| Setting 3 | 0.57 | 0.50 | 0.64 | 0.38 | 0.29 | 0.48 |
| Setting 4 | 0.54 | 0.40 | 0.69 | 0.38 | 0.27 | 0.50 |
| Setting 5 | 0.56 | 0.48 | 0.64 | 0.38 | 0.27 | 0.50 |
| Setting 6 | 0.54 | 0.48 | 0.60 | 0.38 | 0.38 | 0.38 |

For each setting, all datasets were analysed using a semi-parametric Cox PH model (CM, fitted with the `coxph` function of the `R` package `survival`), a semi-parametric promotion time cure model (PTM, fitted with the `PTCMestimBF` function of the `R` package `miCoPTCM`), and a semi-parametric logistic/Cox PH mixture cure model (MCM, fitted with the `smcure` function of the `R` package `smcure`). All models included the treatment covariate, which was included in both parts of the MCM.

### 1.4.2 Simulations results

The results regarding the estimation of the treatment effect in the various settings are presented in Table 1.2; standard errors (s.e.) of the estimated coefficients of the mixture cure model have been estimated using bootstrap with 500 replications. While the estimated coefficients obtained from different models can not be compared together (except for the CM and the PTM) as they do not represent the same quantity, one can compare the estimated cure fractions. They are presented in Table 1.3 and are estimated by $\exp\{-\exp(\boldsymbol{X}^T\hat{\beta})\}$ in the PTM and $1-\hat{\pi}(\boldsymbol{X})$ in the MCM. The *overall* cure rate is estimated, for each dataset, by the average of the estimated individual cure probabilities. Another possibility to compare the fit of the PTM and the MCM in each setting is to consider the estimated conditional survival function of the uncured subjects given by (1.13) for the PTM and by (1.6) for the MCM.

The consequences of a model misspecification can vary largely, depending

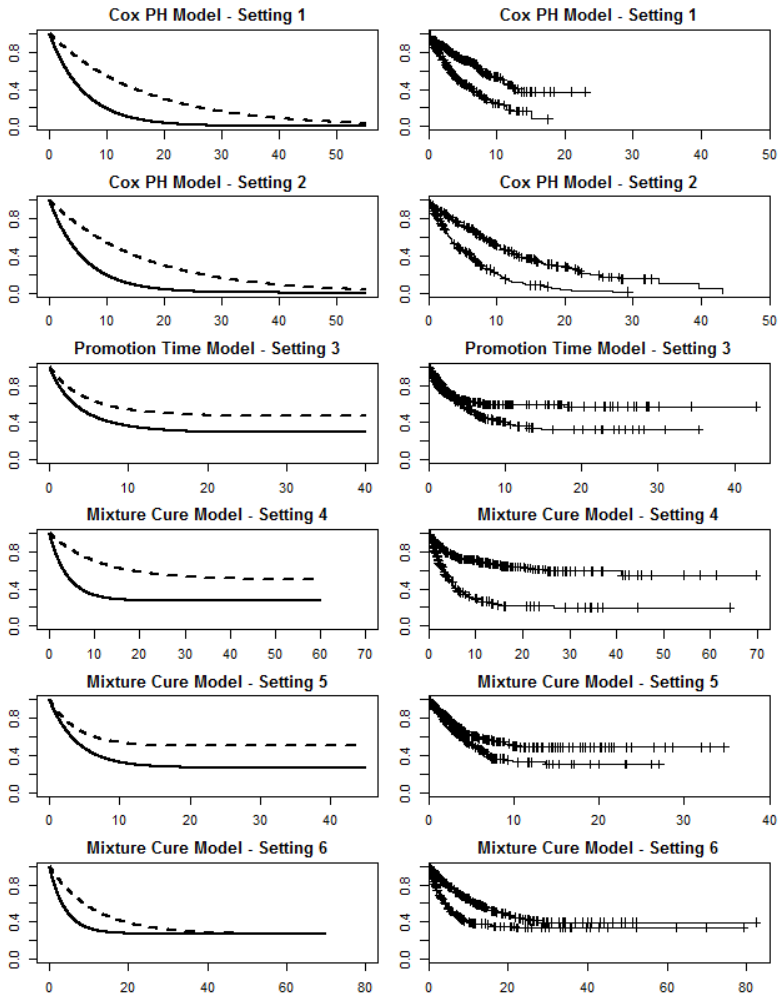



**FIGURE 1.5**
Survival functions in the simulation settings. Left panel: Theoretical survival functions for each simulation setting. Right panel: Estimated survival functions for a random dataset for each setting. The solid line represents the control group ($X = 0$) while the dotted line represents the treatment group ($X = 1$).

**TABLE 1.2**
Simulation results - Estimation of the coefficients. *Note: Emp. S.E.: empirical standard error; Est. S.E.: estimated standard error; Prop. $RH_0$: proportion of cases in which the hypothesis $H_0 : \beta = 0$ was rejected. Empirical standard errors were computed by taking the standard deviation of the estimated values over all replications.*

| | | Cox model | Promotion time model | | Mixture cure model | | |
|---|---|---|---|---|---|---|---|
| | | $\beta_1$ | $\beta_0$ | $\beta_1$ | $\gamma_0$ | $\gamma_1$ | $\beta_1$ |
| **Set. 1** | **True value** | **-1.000** | | | | | |
| | Average | -1.009 | 1.287 | -1.009 | 3.846 | -2.699 | -0.692 |
| | Emp. S.E. | 0.142 | 0.314 | 0.142 | 1.709 | 1.725 | 0.225 |
| | Est. S.E. | 0.140 | 0.190 | 0.139 | 0.454 | 0.552 | 0.198 |
| | Prop. $RH_0$ | 1.000 | 1.000 | 1.000 | 1.000 | 0.971 | 0.779 |
| **Set. 2** | **True value** | **-1.000** | | | | | |
| | Average | -1.004 | 1.999 | -1.004 | 13.504 | -9.822 | -0.926 |
| | Emp. S.E. | 0.113 | 0.218 | 0.113 | 6.936 | 7.118 | 0.124 |
| | Est. S.E. | 0.112 | 0.166 | 0.111 | 3.427 | 3.442 | 0.125 |
| | Prop. $RH_0$ | 1.000 | 1.000 | 1.000 | 0.752 | 0.721 | 1.000 |
| **Set. 3** | **True value** | | **0.200** | **-0.500** | | | |
| | Average | -0.501 | 0.192 | -0.501 | 0.890 | -0.779 | -0.154 |
| | Emp. S.E. | 0.136 | 0.118 | 0.136 | 0.236 | 0.299 | 0.203 |
| | Est. S.E. | 0.139 | 0.110 | 0.139 | 0.199 | 0.302 | 0.193 |
| | Prop. $RH_0$ | 0.960 | 0.408 | 0.960 | 0.998 | 0.745 | 0.133 |
| **Set. 4** | **True value** | | | | **1.000** | **-1.000** | **-1.000** |
| | Average | -1.048 | 0.475 | -1.048 | 1.019 | -0.983 | -1.000 |
| | Emp. S.E. | 0.131 | 0.131 | 0.131 | 0.184 | 0.295 | 0.228 |
| | Est. S.E. | 0.142 | 0.122 | 0.140 | 0.157 | 0.292 | 0.207 |
| | Prop. $RH_0$ | 1.000 | 0.990 | 1.000 | 1.000 | 0.947 | 0.057 |
| **Set. 5** | **True value** | | | | **1.000** | **-1.000** | **0.000** |
| | Average | -0.540 | 0.225 | -0.540 | 1.055 | -1.035 | 0.010 |
| | Emp. S.E. | 0.132 | 0.106 | 0.132 | 0.246 | 0.303 | 0.192 |
| | Est. S.E. | 0.138 | 0.106 | 0.138 | 0.213 | 0.301 | 0.188 |
| | Prop. $RH_0$ | 0.988 | 0.564 | 0.988 | 1.000 | 0.947 | 0.057 |
| **Set. 6** | **True value** | | | | **0.500** | **0.000** | **-1.000** |
| | Average | -0.504 | 0.221 | -0.504 | 0.525 | -0.001 | -0.999 |
| | Emp. S.E. | 0.132 | 0.122 | 0.132 | 0.165 | 0.304 | 0.194 |
| | Est. S.E. | 0.134 | 0.115 | 0.133 | 0.127 | 0.290 | 0.190 |
| | Prop. $RH_0$ | 0.954 | 0.500 | 0.954 | 0.963 | 0.083 | 1.000 |

**TABLE 1.3**
Simulation results - Estimation of the cure rate. *Note: Emp. S.E.: empirical standard error.*

| | | Promotion time model | | | Mixture cure model | | |
|---|---|---|---|---|---|---|---|
| | | Overall | Control | Treatment | Overall | Control | Treatment |
| **Set. 1** | **True value** | **0** | **0** | **0** | **0** | **0** | **0** |
| | Average | 0.155 | 0.038 | 0.273 | 0.150 | 0.043 | 0.258 |
| | Emp. S.E. | 0.140 | 0.030 | 0.103 | 0.132 | 0.038 | 0.102 |
| **Set. 2** | **True value** | **0** | **0** | **0** | **0** | **0** | **0** |
| | Average | 0.038 | 0.001 | 0.074 | 0.029 | 0.003 | 0.055 |
| | Emp. S.E. | 0.046 | 0.002 | 0.040 | 0.041 | 0.006 | 0.044 |
| **Set. 3** | **True value** | **0.380** | **0.295** | **0.477** | **0.380** | **0.295** | **0.477** |
| | Average | 0.389 | 0.298 | 0.480 | 0.383 | 0.293 | 0.473 |
| | Emp. S.E. | 0.100 | 0.042 | 0.041 | 0.101 | 0.048 | 0.048 |
| **Set. 4** | **True value** | **0.380** | **0.269** | **0.500** | **0.380** | **0.269** | **0.500** |
| | Average | 0.385 | 0.202 | 0.568 | 0.379 | 0.267 | 0.491 |
| | Emp. S.E. | 0.187 | 0.041 | 0.042 | 0.123 | 0.035 | 0.062 |
| **Set. 5** | **True value** | **0.380** | **0.269** | **0.500** | **0.380** | **0.269** | **0.500** |
| | Average | 0.384 | 0.286 | 0.481 | 0.378 | 0.261 | 0.495 |
| | Emp. S.E. | 0.105 | 0.038 | 0.040 | 0.125 | 0.046 | 0.045 |
| **Set. 6** | **True value** | **0.380** | **0.378** | **0.378** | **0.380** | **0.378** | **0.378** |
| | Average | 0.379 | 0.288 | 0.470 | 0.373 | 0.373 | 0.374 |
| | Emp. S.E. | 0.101 | 0.043 | 0.043 | 0.051 | 0.038 | 0.061 |

on the true model underlying the data, and on the focus of the estimation: cure probability, conditional survival function, treatment effect size and significance. Due to the link between the Cox PH model and the semi-parametric promotion time cure model with exponential link function (discussed in Section 1.3.2), the results obtained when using these models are clearly very similar. The only difference is that, since the promotion time cure model assumes a cure fraction in the data, the treatment coefficient is interpreted in terms of both its short- and long-term effects.

When there is actually no cure fraction in the data (Settings 1 and 2), the treatment effect is well recovered by the PTM and quite well by the MCM when the censoring is not too high. The estimated coefficients in the incidence part of the MCM are largely biased and accompanied by a very large s.e., showing, as expected, an instability in the estimation of this part of the model. The ability of the models to acknowledge the absence of cure (by estimating a very low cure rate and by appropriately estimating the conditional survival of the uncured patients (data not shown)) is highly dependent on the amount of censoring, as can be seen by comparing the results obtained in Settings 1 and 2. This phenomenon is to be understood in the light of the zero-tail constraint, which is used for identifiability purposes and which treats all censored observations after the last event time as belonging to the cure group: this leads to a positive bias in the estimation of the cure probability, and a negative bias in the estimation of the survival function of the uncured patients.

When there is actually a fraction of cure patients, we have to distinguish situations were the PH assumption may be considered to hold. This is the case when data are generated from a PTM (Setting 3) or from a MCM with a treatment affecting only the incidence (Setting 5, in which the conditional survival functions for both groups level off at the same time point). In that case, it is interesting to note that, although we can not formally compare their coefficients, the PTM and the MCM seem to recover the treatment effect. However, the PTM does not allow us to disentangle the short- and the long-term effects. When the data are generated from a PTM and fitted with a MCM, the true joint treatment effect on both the cure probability and the conditional survival function of the PTM is split into both parts of the model (the average of both estimated coefficients, $\widehat{\beta}_1$ and $\widehat{\gamma}_1$, is incidentally close to the true, unique coefficient). As a result, the significant effect was sometimes recovered for the incidence part, but rarely for the latency. The estimation of the cure rate in each arm as well as of the conditional survival curve for the uncured is nearly unbiased with both the PTM and the MCM. Figure 1.6 displays the results obtained when fitting a MCM on PH data generated by a PTM (Setting 3) and vice-versa (Setting 5).

The situation is however different when data have been generated from a MCM and one can not assume PH anymore, as is the case in Settings 4 and 6. Although, in our simulation setting, the PTM seems to recover some part of treatment effect, the estimated cure rate is biased downwards in the control

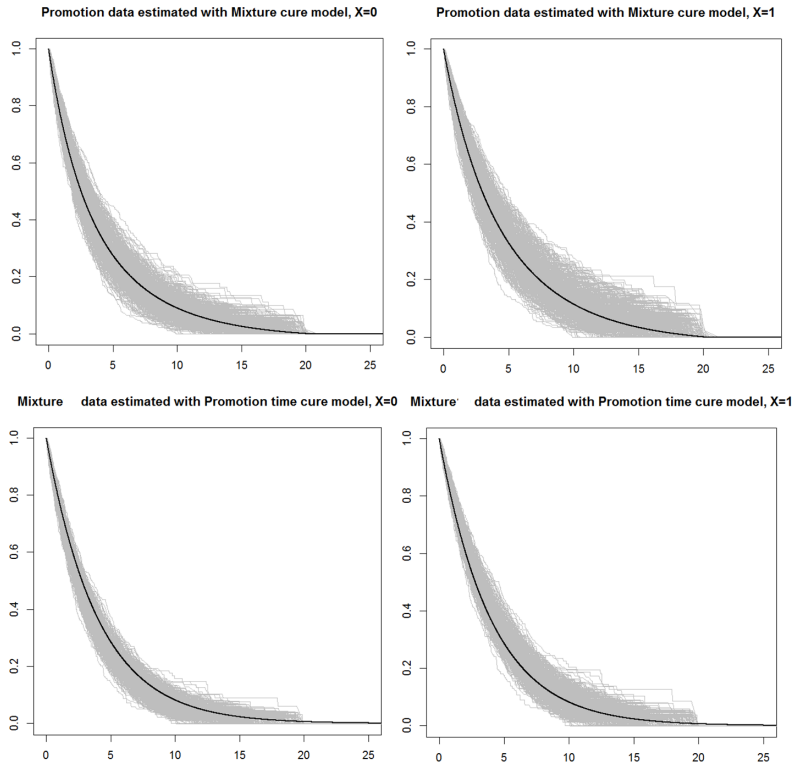



**FIGURE 1.6**
Estimated conditional survival functions for the PTM data estimated with a MCM (top panel) and for the MCM data estimated with a PTM (bottom panel), for the control arm (left panel) and the treatment arm (right panel).

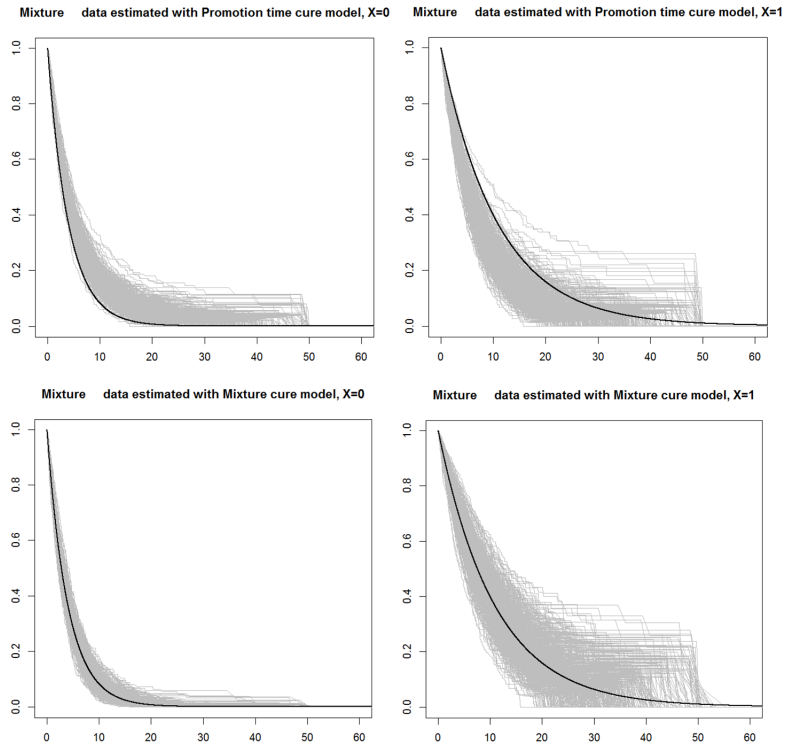



**FIGURE 1.7**
Estimated conditional survival functions for the MCM data (setting 4) estimated with a PTM (top panel) and with a MCM (bottom panel), for the control arm (left panel) and the treatment arm (right panel).

arm and upward in the treatment arm. As shown in the top of Figure 1.7, this leads to an overestimation of the conditional survival in the control group and an underestimation in the treatment group. In the control arm, the PTM assigns too few patients to the cure group and hence estimates a too high survival for the uncured; the opposite holds in the treatment group (too many patients to the cure group and a too low survival for the uncured). These biases are due to the model misspecification; only one coefficient is estimated, which tries to summarize both effects (on the incidence and the latency). As a result, none of the two effects is correctly estimated. As expected, there is no such problems when estimating these data with the appropriate model, see bottom of Figure 1.7.

### 1.4.3 Discussion of the results and recommendations

Cure models are still rarely used in cancer clinical trials, despite the fact that there are nowadays several cancer types for which we can expect a fraction of the population to be cured. One argument against the use of the cure models is that as long as the PH assumption is met, the Cox PH model provides reliable estimates of the treatment effect [36]. Indeed, our simulations show that if the cure fraction is ignored and a Cox PH model is fitted, the treatment coefficient is perfectly recovered in size and significance in cases where we have PH. This is indeed true and actually not surprising, given the mathematical link between the Cox PH model and the semi-parametric promotion time cure model [42]. In the specific context of our simulations, when the PH assumption does not hold, the Cox PH model also allows one to recover a significant treatment effect, whose estimated value appears to be close to the average of both true coefficients. However, using the promotion time cure model instead of the Cox PH model has the advantage of making it clear that the coefficient associated with the treatment should be interpreted both in terms of short- and long-term effects.

One may argue that, unless the disease is always fatal, the proportion of patients being cured should be considered as an important component of the survival benefit, rather than just considering the HR or median time to failure [32]. If the PH assumption is met, both the Cox PH model and the promotion time cure model can actually be used to obtain an estimate of the cure fraction, which is clearly a useful piece of information in the evaluation of curative treatment. However, one has to be careful that this proportion may be overestimated by the promotion time cure model if there is actually no cure fraction and if the censoring is high. So, like many other authors (see for example [47]), we recommend not using such models when there is no evidence of cure.

Furthermore, the Cox PH and the promotion time models should not be used whenever the PH assumption is not met. If the reason of this non-proportionality is the presence of a cure fraction, which can be assessed from the presence of a long plateau including a sufficient number of censored observations, then the mixture cure model is a flexible alternative to be considered. Both parametric and semi-parametric versions are easily accessible in `R` and `SAS`, and both Cox PH and AFT models can be considered in the latency part. The (semi-)parametric logistic/Cox PH model provides parameter estimates which have an easy interpretation and can be translated into ORs or RRs for the probability of cure and in HRs for the event-times amongst the uncured, providing a clear view on the short- and long-term effects of the treatment.

The statement that "*as long as one can assume that not all patients will experience the event of interest, a cure model should be preferred*" needs to be nuanced. First, assuming that there is a cure fraction is not enough; we must have evidence of it, through sufficient follow-up, in order for a cure model to perform well. Second, even when there is such evidence from the data,

if we still have PH and if we are not particularly interested in splitting the curative effect from an event-delaying effect, nor to emphasize the presence of a cure fraction, then the classical Cox PH model can indeed still be used. However, whenever there exists a fraction of cured or long-term survivors, additional information (or, even, more correct information, in case of non-PH) can be gained from using an appropriate cure model analysis compared with a standard Cox analysis. Unfortunately, there are up to now no clear criteria on what is evidence of a cure fraction and no widely available statistical way to test whether there is "a sufficiently long plateau containing enough censored observations". Some attempts to develop statistical tests on the presence of cure have been proposed [27, 33, 9, 63, 21] but they have not been implemented in available software. As a consequence, one has mainly to rely on a visual inspection of the tail of the KM estimated survival curves.

A key ingredient in clinical trials is the design phase, and in particular the sample size calculation. All standard procedures for sample size calculation with time-to-event endpoints actually rely on the PH assumption and can therefore not be used if we expect the presence of a cure fraction that could put this assumption in jeopardy. A sample size formula for the logistic/PH mixture cure model has been proposed by [55] and later implemented in the `R` package `NPHMC` [10]. This formula can be used to compute the required sample size for testing differences in the short- and/or long-term outcome of the patients, and can account for various accrual patterns. Furthermore, the `NPHMC` package allows one to choose for the latency part of the model between a parametric PH model (exponential or Weibull) or a Cox semi-parametric PH model. Numerical examples and simulations results are presented and show that ignoring the cure rate can lead to either underpowered or overpowered studies [55].

## 1.5 Melanoma clinical trial

The ECOG phase III clinical trial e1684 was set up to evaluate the effect on relapse free survival (RFS) of high dose interferon alpha-2b (IFN) versus placebo (PBO) as postoperative adjuvant therapy. A total of $n = 285$ patients were randomized to either IFN ($n = 145$) or PBO ($n = 140$). Two additional covariates are included in the freely available database: gender (39.8% female) and age (centered to the mean). The main analysis of this trial, as it appears in the original publication of the trial results, does not take cure into account [24]. However, there is clear evidence, both from a medical point of view and by inspecting the estimated survival curves in each treatment group (see Figure 1.8, left panel), of a presence of a cure fraction. These data have already been extensively used to illustrate publications on cure models, see for example [15, 55, 22].

The dataset is freely available in the R package `smcure` and can be loaded in R using the following command:

```
library(smcure)
data(e1684)
```

The estimated KM survival curve for RFS shows a plateau starting at around 6 years, with however an event still occurring at around 8 years in the control group. Both estimated survival curves run in parallel and reach their plateau at about the same time; in line with Section 1.4, we will consider that the results from both the PTM and the MCM can be trusted.

A classical semi-parametric Cox PH model (CM) can be fitted on these data with the `R` package `survival`, as follows:

```
library(survival)
cox <- coxph(Surv(FAILTIME,FAILCENS==1)~TRT+AGE+SEX,e1684)
summary(cox)
```

A semi-parametric promotion time model (PTM) with exponential link function can be fitted with the `R` package `miCoPTCM`:

```
library(miCoPTCM)
vc <- matrix(nrow=4,ncol=4,0)
ptcm <- PTCMestimBF(formula=Surv(e1684$FAILTIME,e1684$FAILCENS)
                    ~TRT+SEX+AGE,data=e1684,varCov=vc,
                    init=runif(4))
summary(ptcm)
```

Finally, a semi-parametric logistic/Cox PH mixture cure model (MCM) can be estimated with the `R` package `smcure`:

```
mcm <- smcure(Surv(FAILTIME,FAILCENS)~TRT+SEX+AGE,
              cureform=~TRT+SEX+AGE,data=e1684,model="ph",
              nboot=500)
printsmcure(mcm)
```

Results (obtained with the three models) for the treatment effect, adjusted for gender and sex, are displayed in Table 1.4.

As expected, the results obtained with CM and PTM are similar, with the interpretation of the $\beta$ coefficients being linked to both short- and long-term effects. To interpret these results, we have to keep in mind that IFN is coded 1 while PBO is coded 0. The treatment effect, adjusted for gender and age, is $-0.365$ for the PTM model, indicating an advantage for patients treated with IFN ($p < 0.05$). The MCM model allows us to disentangle the effect of IFN on the occurrence and on the timing of the event. The OR for the probability to be uncured, adjusted for gender and age, is $\exp(-0.588) = 0.556$ corresponding to a lower risk of being uncured in the IFN group, and thus a higher risk to be cured in this group. Regarding the latency part of the model, the HR for

the uncured patients is $\exp(-0.154) = 0.857$ in favor of the IFN group too. It appears from these results that IFN indeed has a beneficial effect on the RFS of the patients but that this effect is mainly a long-term one, acting on the probability to be cured. The fact that this OR is not significant might be explained by the fact that twice the number of parameters have to be estimated in the MCM compared to the PTM. For a male ($x_{male} = 0$) with average age ($x_{age} = 0$), the estimated cure fraction can be recovered from the PTM as

$$\begin{aligned} \text{PBO arm} &: \exp(-\exp(0.412)) = 0.221 \\ \text{IFN arm} &: \exp(-\exp(0.412 - 0.365)) = 0.351 \end{aligned}$$

and about the same values can be retrieved from the MCM:

$$\begin{aligned} \text{PBO arm} &: 1 - \frac{\exp(1.365)}{1 + \exp(1.365)} = 0.203 \\ \text{IFN arm} &: 1 - \frac{\exp(1.365 - 0.588)}{1 + \exp(1.365 - 0.588)} = 0.315 \end{aligned}$$

The estimated population RFS curves by treatment arm from the PTM and the MCM for male patients with average age are displayed on the right panel of Figure 1.8. Curves obtained from both models are very similar and we find back at the tails of these curves the results given above about cure rate estimation.

**TABLE 1.4**
Melanoma data: results from semi-parametric Cox PH model (CM), semi-parametric logistic/Cox PH model (MCM), and semi-parametric promotion time cure model (PTM) for treatment (0: control and 1: treatment) adjusted for age (0: male and 1: female) and gender (centered to the mean). *Note: S.E.: standard error.*

| | | CM | MCM | | PTM |
|---|---|---|---|---|---|
| | | | Incidence | Latency | |
| Intercept | Estimate | | 1.365 | | 0.412 |
| | S.E. | | 0.322 | | 0.139 |
| | P-value | | 0.000 | | 0.003 |
| Treatment | Estimate | -0.360 | -0.588 | -0.154 | -0.365 |
| | S.E. | 0.144 | 0.349 | 0.169 | 0.154 |
| | P-value | 0.012 | 0.092 | 0.363 | 0.017 |
| Age | Estimate | 0.005 | 0.020 | -0.008 | 0.005 |
| | S.E. | 0.005 | 0.016 | 0.006 | 0.005 |
| | P-value | 0.357 | 0.205 | 0.212 | 0.358 |
| Gender | Estimate | -0.018 | -0.087 | 0.099 | -0.018 |
| | S.E. | 0.147 | 0.328 | 0.172 | 0.159 |
| | P-value | 0.903 | 0.791 | 0.564 | 0.909 |

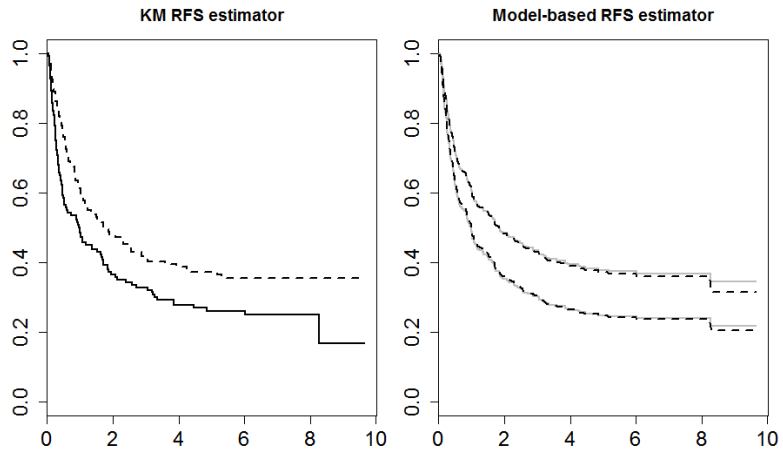



**FIGURE 1.8**
Melanoma data. Left panel: Kaplan-Meier estimated population RFS curves (the solid line represents the IFM arm and the dotted line the PBO arm). Right panel: Model-based estimated population RFS for male patients of average age as obtained from the PTM model (gray solid lines) and from the MCM (black dotted lines).

# Bibliography


[1] M. Amico and I. Van Keilegom. Cure models in survival analysis. *Annual Review of Statistics and Its Application*, 5(1), 2018.

[2] M. Amico, I. Van Keilegom, and C. Legrand. The single-index / Cox mixture cure model. *Under revision*, 2018.

[3] T. M.-L. Andersson and P. C. Lambert. Fitting and modeling cure in population-based cancer studies within the framework of flexible parametric survival models. *The Stata Journal*, 12:623–638, 2012.

[4] J. Berkson and R. P. Gage. Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, 47:501–515, 1952.

[5] W.A. Bleyer. Acute lymphoblastic leukaemia in children: advances and prospectus. *Cancer*, 65:689–695, 1990.

[6] J. W. Boag. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society, Series B*, 11:15–44, 1949.

[7] V. Bremhorst, M. Kreyenfeld, and P. Lambert. Fertility progression in Germany: An analysis using flexible nonparametric cure survival models. *Demographic Research*, 35(18):505–534, 2016.

[8] V. Bremhorst and P. Lambert. Flexible estimation in cure survival models using Bayesian P-splines. *Computational Statistics and Data Analysis*, 93:270–284, 2016.

[9] P. Broet, Y. De Rycke, P. Tubert-Bitter, J. Lellouch, B. Asselain, and T. Moreau. A semiparametric approach for the two-sample comparison of survival times with long-term survivors. *Biometrics*, 57:844–852, 2001.

[10] C. Cai, S. Wang, W. Lu, and J. Zhang. Nphmc: An R-package for estimating sample size of proportional hazards mixture cure models. *Computer Methods and Programs in Biomedicine*, 113(1):290–300, 2014.

[11] C. Cai, Y. Zou, Y. Peng, and J. Zhang. Smcure: An R-package for estimating semiparametric mixture cure models. *Computer Methods and Programs in Biomedicine*, 108(3):1255–1260, 2012.

[12] M.-H. Chen, J. G. Ibrahim, and D. Sinha. A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, 94:909–919, 1999.

[13] F. Cooner, S. Banerjee, B. Carlin, and D. Sinha. Flexible cure rate modeling under latent activation scheme. *Journal of the American Statistical Association*, 102:560–572, 2007.

[14] F. Corbière, D. Commenges, J. M. G. Taylor, and P. Joly. A penalized likelihood approach for mixture cure models. *Statistics in Medicine*, 28:510–524, 2009.

[15] F. Corbière and P. Joly. A SAS macro for parametric and semiparametric mixture cure models. *Computer Methods and Programs in Biomedicine*, 85(2):173–80, 2007.

[16] D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society, Series B*, 34:187–220, 1972.

[17] V. T. Farewell. A model for a binary variable with time-censored observations. *Biometrika*, 64:43–46, 1977.

[18] V. T. Farewell. The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, 38:1041–1046, 1982.

[19] M.E. Ghitany, R.A. Maller, and S. Zhou. Exponential mixture models with long term survivors and covariates. *Journal of Multivariate Analysis*, 49:218–241, 1994.

[20] Y. Gu, D. Sinha, and S. Banerjee. Analysis of cure rate survival data under proportional odds model. *Lifetime Data Analysis*, 17:123–134, 2011.

[21] W.W. Hsu, D. Todem, and K.M. Kim. A sup-score test for the cure fraction in mixture cure models for long-terms survivors. *Biometrics*, 76:1348–1357, 2016.

[22] J. G. Ibrahim, M.-H. Chen, and D. Sinha. Bayesian semiparametric models for survival data with a cure fraction. *Biometrics*, 57:383–388, 2001.

[23] J. G. Ibrahim, M.-H. Chen, and D. Sinha. *Bayesian Survival Analysis*. Springer-Verlag, Boca Raton, 2001.

[24] J.M. Kirkwood, M.H. Strawderman, M.S. Ernstoff, T.J. Smith, E.C. Borden, and Blum R.H. Interefron alfa-2b adjuvant therapy of high-risk resected cutaneous melanoma: the Eastern Cooperative Oncology Group Trial EST 1684. *Journal of Clinical Oncology*, 14:7–17, 1996.

[25] A. Y. C. Kuk and C.-H. Chen. A mixture model combining logistic regression with proportional hazards regression. *Biometrika*, 79:531–541, 1992.

[26] P. C. Lambert. Modeling of the cure fraction in survival studies. *The Stata Journal*, 7:1–25, 2007.

[27] E.M. Laska and M.J. Meisner. Nonparametric estimation and testing in a cure model. *Biometrics*, 48:1223–1234, 19992.

[28] C.-S.. Li and J. M. G. Taylor. A semi-parametric accelerated failure time cure model. *Statistics in Medicine*, 21:3235–3247, 2002.

[29] H. Liu and Y. Shen. A semiparametric regression cure model for interval-censored data. *Journal of the American Statistical Association*, 104:1168–1178, 2009.

[30] W. Lu. Maximum likelihood estimation in the proportional hazards cure model. *Annals of the Institute of Statistical Mathematics*, 60:545–574, 2008.

[31] Y. Ma and G. Yin. Cure rate model with mismeasured covariates under transformation. *Journal of the American Statistical Association*, 103:743–756, 2008.

[32] S. Maetani and W. Gamel. Parametric cure model versus proportional hazards model in survival analysis of breast cancer and other malignancies. *Advances in Breast Cancer research*, 2:119–15, 2013.

[33] R.A. Maller and S. Zhou. *Survival Analysis With Long Term Survivors*. Wiley, New York, 1996.

[34] N. Mantel. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*, 50:163–70, 1966.

[35] M. Othus, B. Barlogie, M. L. LeBlanc, and J. J. Crowley. Cure models as a useful statistical tool for analyzing survival. *Statistics in Clinical Cancer Research*, 18:3731–3736, 2012.

[36] X. Paoletti and B. Asselain. Survival analysis in clinical trials: Old tools or new techniques. *Surgical Oncology*, 19:55–58, 2010.

[37] V. Patilea and I. Van Keilegom. A general approach for cure models in survival analysis. *Submitted*, 2018.

[38] Y. Peng and K. B. G. Dear. A nonparametric mixture model for cure rate estimation. *Biometrics*, 56:237–243, 2000.

[39] Y. Peng, K. B. G. Dear, and J. W. Denham. A generalized $f$ mixture model for cure rate estimation. *Statistics in Medicine*, 17:813–830, 1998.

[40] Y. Peng and J.M.G. Taylor. Cure models. In *Handbook of Survival Analysis*, pages 113–134. In Klein, J., van Houwelingen, H., Ibrahim, J. G., and Scheike, T. H., editors. Chapman and Hall, Boca Raton, FL, USA, 2014.

[41] F. Portier, A. El Ghouch, and I. Van Keilegom. Efficiency and bootstrap in the promotion time cure model. *Bernoulli*, 23:3437–3468, 2017.

[42] F. Portier, A. El Ghouch, and I. Van Keilegom. On proportional hazards cure models. *Submitted*, 2017.

[43] L.E. Rutqvist, A. Wallgren, and B. Nilsson. Is breast cancer a curable disease? A study of 14,731 women with breast cancer from the Cancer Registry of Norway. *Cancer*, 53:1793–1800, 1984.

[44] D. Sargent, A. Sobrero, A. Grothey, M.J. OConnell, M. Buyse, et al. Evidence for cure by adjuvant therapy in colon cancer: Observations based on individual patient data from 20, 898 patients on 18 randomized trials. *Journal of Clinical Oncology*, 27:812–877, 2009.

[45] S. Scolas, A. El Ghouch, C. Legrand, and A. Oulhaj. Variable selection in a flexible parametric mixture cure model with interval-censored data. *Statistics in medicine*, 35:1210–1225, 2016.

[46] R. Sposto. Cure model analysis in cancer: an application to data from the Children's Cancer Group. *Statistics in Medicine*, 21:293–312, 2002.

[47] J. P. Sy and J. M. G. Taylor. Estimation in a Cox promotional hazards cure model. *Biometrics*, 56:227–236, 2000.

[48] J. M. G. Taylor. Semi-parametric estimation in failure time mixture models. *Biometrics*, 51:899–907, 1995.

[49] L. Thomas and E.M. Reyes. Tutorial: Survival estimation for Cox regression models with time-varying coefficients using SAS and R. *Journal of Statistical Software*, 61:1–23, 2014.

[50] M. Tournoud and R. Ecochard. Promotion time models with time-changing exposure and heterogeneity: Application to infectious diseases. *Biometrical Journal*, 50:395–407, 2008.

[51] A. Tsodikov. A proportional hazards model taking account of long-term survivors. *Biometrics*, 54:1508–1516, 1998.

[52] A. Tsodikov. 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*, 21:895–920, 2002.

[53] A. Tsodikov, J. G. Ibrahim, and A. Y. Yakovlev. Estimating cure rates from survival data: An alternative to two-component mixture models. *Journal of the American Statistical Association*, 98:1063–1078, 2003.

[54] J. Wang and S. K. Ghosh. Shape restricted nonparametric regression with Bernstein polynomials. *Computational Statistics and Data Analysis*, 56:2729–2741, 2012.

[55] S. Wang, J. Zhang, and W. Lu. Sample size calculations for the proportional hazards cure model. *Statistics in Medicine*, 31(29):3959–3971, 2012.

[56] A. Y. Yakovlev, B. Asselai, V.-J. Bardou, A. Fourquet, T. Hoang, A. Rochefedière, and A. D. Tsodikov. A simple stochastic model of tumor recurrence and its application to data on premenopausal breast cancer. In B. Asselain, M. Boniface, C. Duby, C. Lopez, J.-P. Masson, and J. Tranchefort, editors, *Biométrie et Analyse de Données Spatio-Temporelles*, pages 66–82. Société française de biométrie, ENSA Rennes, 1993.

[57] A. Y. Yakovlev and A. D. Tsodikov. *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific, Singapore, 1996.

[58] K. Yamaguchi. Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of permanent employment in Japan. *Journal of the American Statistical Association*, 87(418):284–282, 1992.

[59] Y. Yilmaz, J.F. Lawless, I.L. Andrulis, and S.B. Bull. Insight from mixture cure modeling of molecular markers for prognosis in breast cancer. *Journal of Clinical Oncology*, 31(16):2047–2054, 2013.

[60] G. Yin and J. G. Ibrahim. Cure rate models: a unified approach. *Canadian Journal of Statistics*, 33:559–570, 2005.

[61] D. Zeng, G. Yin, and J. G. Ibrahim. Semiparametric transformation models for survival data with a cure fraction. *Journal of the American Statistical Association*, 101:670–684, 2006.

[62] J. Zhang and Y. Peng. A new estimation method for the semiparametric accelerated failure time mixture cure model. *Statistics in Medicine*, 26:3157–3171, 2007.

[63] Y. Zhao, A.H. Lee, K.K.W. Yau, V. Burke, and G.J. McLachlan. A score test for assessing the cured proportion in long-term survivor mixture cure model. *Statistics in Medicine*, 28:3454–3466, 2009.