CHAPTER 3

Assessment of predictive accuracy of an intermittently observed binary time-dependent marker

This chapter is based on joint work with Hein Putter and Marta Fiocco.

Abstract

Following tumor removal surgery soft tissue sarcoma patients are at risk for disease recurrence, which can indicate an increased risk of death. The predictive value of this time-dependent variable can be summarized by the time-specific Area Under the receiver operating characteristics Curve (AUC). However, the fact that recurrence is often diagnosed in an interval-censored fashion is frequently ignored when modelling its effect on survival. Follow-up schemes determine the times at which a patient is diagnosed with recurrence. The effect that ignoring the interval-censored nature of the observation time has on the time-specific AUC in both incident/dynamic and cumulative/dynamic definition is studied.[80, 171] AUC estimates derived from different methods for fitting two types of models are compared: the Cox model with time-dependent covariate and the illness-death model for interval-censored data. Data is simulated from an illness-death model with Weibull transition hazards and the disease state is censored at regular observation intervals. The true AUC is determined by transition probabilities, derived from the Weibull transition hazards. The method is applied to a data set of 2232 patients with high-grade soft tissue sarcoma and results are discussed.

§3.1 Introduction

Survival analysis studies the distribution of time from a time origin to an event of interest. It is often applied in the medical field where for example the time from diagnosis to death is studied. The intrinsic particularity of survival data is that it is generally incomplete: the event of interest cannot always be observed because it takes time to observe it. Data of individuals who did not experience the event of interest within a specific time window are right-censored. A frequently used method to study the effect of covariates on survival time is the Cox proportional hazards model.[44] In the medical field it is often applied to study the effect of risk factors on a single event such as death or disease progression. However, in practice disease progression may be described by more than one type of event. These more complicated event structures can be modeled simultaneously using multi-state models.[119] The most simple of such models is the illness-death model, which is described by three states (see Figure 3.1): an individual is initially disease-free (state 0), he may then develop disease (state 1) and die (state 2) or he may die without disease. Like in the single event situation the Cox model can be used to model the effect of covariates on the transitions between states.

The illness-death model is applicable to a variety of disease settings; a problem arises, however, if the time of disease cannot be observed exactly. Often, disease can only be diagnosed at pre-specified follow-up times. An example lies in the care of patients with soft tissue sarcoma. After initial treatment by tumor removal surgery a patient may develop distant metastases and then die. Metastases are diagnosed at pre-specified follow-up visits at which an X-ray of the patient is screened. If metastases are found, it is therefore only known that they appeared between the last negative screening and the first positive screening; the data is interval-censored. This type of data contains two types of missing information: (1) the time of disease is only known to have happened between two visits, it is interval-censored. (2) If the last disease screening prior to death or last recorded follow-up was negative the disease status of a patient between last screening and death or last recorded follow-up is unknown.

The illness-death model for interval-censored data has been previously studied and it was found that ignoring the observation scheme of the data leads to biased estimates of regression coefficients, baseline hazards, and survival.[64, 86, 65, 168, 97] A prominent motivation comes from the study of dementia data.[86, 168, 97] Dementia is diagnosed at infrequent follow-up visits which results in the time to dementia being interval-censored. Further, if a patient's last dementia test was negative and he dies it is not known if he acquired dementia prior to death. Frydman (1995)[64] developed a non-parametric maximum likelihood procedure for the estimation of the cumulative transition hazards when times of disease are interval-censored. He does not address the second form of incompleteness however, i.e. it is assumed that the disease state is known before death or right-censoring time. Joly et al. (2002)[86] proposed a non-parametric penalized likelihood method to estimate transition intensities in an illness-death model with an intermittently observed disease state. Simulations showed that not adjusting for the interval-censored nature of the data leads to a systematic bias in the estimation of transition intensities. Frydman and Szarek (2009)[65] ex-

tended the methodology of Frydman (1995)[64] to incorporate the observations with unknown intermediate event status. They estimated the distribution of the time to the first occurrence of disease or death and showed that their method corrects bias. Yu et al. (2010)[168] used multiple imputation to analyze two aspects concerning the risk of dementia: the risk of developing dementia and the impact of dementia on survival. Leffondré et al. (2013)[97] performed simulation studies to show how interval-censoring affects the estimation of the effect of risk factors.

If event times are observed exactly the illness-death model can be estimated with several R-packages, such as the `survival` and the `mstate` package.[141, 49, 48] The number of packages that can deal with an interval-censored disease state however is limited. The `msm` and the `SmoothHazard` package can fit an illness-death model for interval-censored disease times and exact death times.[85, 144] In the `msm` package piece-wise constant hazards need to be assumed and in the `SmoothHazard` package the user is able to choose between Weibull transition hazards and M-splines. The `coxinterval` package can estimate the illness-death model for data with interval-censored disease times as long as some disease times are observed exactly.[31] While the effect of ignoring the interval-censored nature of the data on regression coefficients and baseline hazards has been studied, the effect on the assessment of predictive accuracy has been neglected so far.

The aim of this article is to study the predictive accuracy of an interval-censored binary disease marker on survival. How much does the occurrence or absence of disease contribute to survival predictions over time? The illness-death model for data in which the disease state is interval-censored is considered. The effect of interval-censoring on the time-specific Area Under the receiver operating characteristics Curve (AUC) in both incident/dynamic and cumulative/dynamic definition is evaluated.[80, 171] Several estimation approaches are compared for two types of models: the Cox model with time-dependent disease marker and the illness-death model for interval-censored data as implemented in the `msm` and `SmoothHazard` R-packages.[85, 144] For this purpose a simulation study is conducted where data is simulated from an illness-death model with Weibull transition hazards.

The remainder of this article is organized as follows. Section 3.2 introduces the definitions of time-specific AUC for a binary time-dependent marker and the theoretical AUC values for a Weibull illness-death model. In Section 3.3 the different models considered in this work are illustrated. A simulation study is presented in Section 3.4. In Section 3.5 the different methods are applied to data of soft tissue sarcoma patients. A discussion follows in Section 3.6.

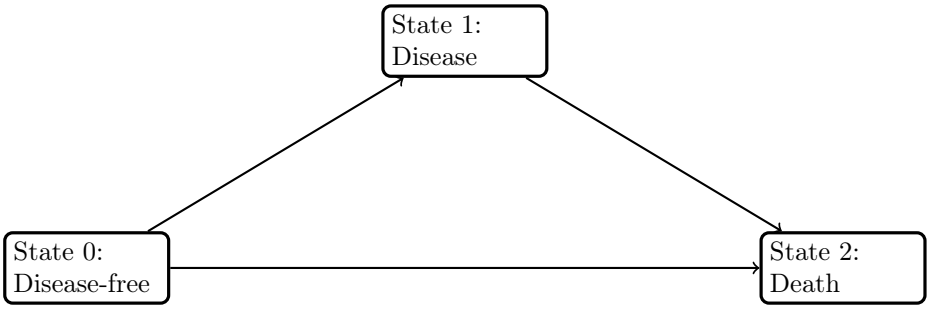

Figure 3.1: Illness-death model.

§3.2 Time-specific AUC for binary marker

Several measures of predictive accuracy have been introduced in the field of survival analysis. In this article predictive accuracy is assessed using the time-specific definitions of sensitivity and specificity which allow for censoring proposed by Heagerty et al. (2000)[79], Heagerty and Zheng (2005)[80], and Zheng and Heagerty (2007)[171].

Originally, sensitivity and specificity were defined considering a binary outcome B. Individuals with outcome $B = 1$ were considered to be 'cases' and individuals with outcome $B = 0$ were considered 'controls'. A covariate X together with a classification criterion c can then be used as a classification rule: a subject is predicted to be a 'case' if the value of the covariate is bigger than c and it is predicted to be a 'control' otherwise. The accuracy of this classification rule can be summarized by the correct classification rates; sensitivity$(c) = P(X > c|B = 1)$ and specificity$(c) = P(X \leq c|B = 0)$. The full range of sensitivity and specificity for different classification criteria c can be graphically summarized by the Receiver Operation Characteristic (ROC) curve which plots sensitivity against 1-specificity. The ROC curve illustrates the difference of the marker distribution between cases and controls. If the distributions are the same, which means that the marker is useless to distinguish cases from controls, then the ROC curve lies on the 45 degree line. The Area Under the Curve (AUC) is a measure of concordance between the marker and the outcome and can be used to summarize the predictive accuracy of the marker X. It is defined by

$$\text{AUC}(X) = P(X_1 > X_0) + 0.5 \cdot P(X_1 = X_0),$$

where X_1 is the value of a covariate drawn from the distribution of cases ($B = 1$) and X_0 is the value of a covariate drawn from the distribution of controls ($B = 0$). To extend the concept of sensitivity and specificity to allow for censored data several definitions for cases and controls were studied[79, 80, 171].

In this article a time-dependent binary covariate $X(t)$ representing disease is considered. The covariate $X(t)$ can take values 0 and 1 which correspond to not having disease and having disease at time t, respectively. The Markov assumption is assumed for the studied illness-death model throughout the article.

§3.2.1 Incident cases and dynamic controls

Heagerty and Zheng (2005)[80] define incident sensitivity and dynamic specificity at time t as

$$\text{sensitivity}^I(c,t) = P(X(t) > c \mid T = t),$$
$$\text{specificity}^D(c,t) = P(X(t) \leq c \mid T > t),$$

where c is a classification criterion, T is time of death and $X(t)$ is the time-dependent disease marker evaluated at time t. In this definition the individuals who die at time t are considered cases and individuals who survive beyond time t are considered controls. Let i, j be individuals, $X_i(t), X_j(t)$ their marker values at time t, and T_i and T_j their death times. The incident/dynamic AUC is then defined by[80]

$$\begin{aligned}\text{AUC}^{I/D}(t) =& P(X_i(t) > X_j(t) \mid T_i = t, T_j > t) \\ &+ 0.5P(X_i(t) = X_j(t) \mid T_i = t, T_j > t).\end{aligned}$$

In case $X_i(t)$ and $X_j(t)$ are binary covariates the $\text{AUC}^{I/D}(t)$ can be rewritten as

$$\text{AUC}^{I/D}(t) = 0.5 + 0.5(p(t) - \pi_1(t)), \tag{3.2.1}$$

where $\pi_1(t)$ is the probability that a person alive at time t has experienced disease (prevalence of disease) and $p(t)$ is the probability that a person who dies at time t has a history of disease (see Appendix 3.A). The disease marker $X(t)$ is related to the illness-death model of Figure 3.1 in the following way: $X(t) = 0$ if a patient did not move to state 1 (disease) before time t (in state 0 or 2 at time t) and $X(t) = 1$ if a patient moved to state 1 (disease) before time t (in state 1 or 2 at time t). The terms $\pi_1(t)$ and $p(t)$ can be expressed by transition probabilities in a multi-state model with states 0, 1, 2 (Figure 3.1),

$$\pi_1(t) = P(X_i(t) = 1 \mid T_i > t) = \frac{P_{01}(t)}{P_{00}(t) + P_{01}(t)}, \tag{3.2.2}$$

$$p(t) = P(X_i(t-) = 1 \mid T_i = t) = \frac{\frac{\lambda_{12}(t)}{\lambda_{02}(t)} P_{01}(t-)}{P_{00}(t-) + \frac{\lambda_{12}(t)}{\lambda_{02}(t)} P_{01}(t-)}, \tag{3.2.3}$$

where $t-$ means just before time t, $P_{0l}(t)$ is the conditional probability of being in state l, $(l = 0, 1)$ at time t given in state 0 at time 0 and $\lambda_{k2}(t)$ is the transition hazard at time t for moving from state k, $(k = 0, 1)$ to state 2.

The incident/dynamic AUC at a specific time t measures how well the disease marker evaluated at time t separates those who die at t from those who survive.

The difference between $p(t)$ and $\pi_1(t)$ is equal to

$$\begin{aligned} p(t) - \pi_1(t) =& \frac{\gamma(t)P_{01}(t)}{P_{00}(t) + \gamma(t)P_{01}(t)} - \frac{P_{01}(t)}{P_{00}(t) + P_{01}(t)} \\ =& (\gamma(t) - 1)\frac{P_{01}(t)P_{00}(t)}{(P_{00}(t) + \gamma(t)P_{01}(t))(P_{00}(t) + P_{01}(t))}, \\ =& (\gamma(t) - 1)\frac{1}{(1 + \gamma(t)P_{01}(t)/P_{00}(t))(1 + P_{00}(t)/P_{01}(t))}, \end{aligned} \tag{3.2.4}$$

where $\gamma(t) = \frac{\lambda_{12}(t)}{\lambda_{02}(t)}$.

From (3.2.1) and (3.2.4) follows that if $\gamma(t) \equiv 1$ then $AUC^{I/D}(t) = 0.5$, if $\gamma(t) > 1$ then $AUC^{I/D}(t) \geq 0.5$ and if $\gamma(t) < 1$ then $AUC^{I/D}(t) \leq 0.5$.

§3.2.2 Cumulative cases and dynamic controls

Zheng and Heagerty (2007)[171] define cumulative sensitivity and dynamic specificity at time t for a time-dependent covariate evaluated at time s as

$$\begin{aligned} \text{sensitivity}^C(c \mid \text{start} = s, \text{stop} = t) &= P(X(s) > c \mid T \geq s, T \leq t), \\ \text{specificity}^D(c \mid \text{start} = s, \text{stop} = t) &= P(X(s) \leq c \mid T \geq s, T > t), \end{aligned}$$

where T is time of death, $X(s)$ is marker measurement at time s. Cases are individuals who die within a time window $(t - s)$ from s and controls are individuals who survive the time window. The cumulative/dynamic AUC is then defined by

$$\begin{aligned} \text{AUC}^{C/D}(s,t) =& P(X_i(s) > X_j(s) \mid T_i > s, T_i \leq t, T_j > s, T_j > t) \\ &+ 0.5P(X_i(s) = X_j(s) \mid T_i > s, T_i \leq t, T_j > s, T_j > t) \end{aligned}$$

where i, j are individuals, $X_i(s), X_j(s)$ their marker values at time s, and T_i, T_j their death times. For binary $X_i(s)$ and $X_j(s)$ the $\text{AUC}^{C/D}$ can be rewritten as

$$\text{AUC}^{C/D}(s,t) = 0.5 + 0.5(p(s,t) - \pi_1(s,t)), \tag{3.2.5}$$

where $\pi_1(s,t)$ is the probability that a person alive at time t had experienced disease by time s and $p(s,t)$ is the probability that a person that dies in the time interval $(s,t]$ had experienced disease by time s (see Appendix 3.A). The quantities $\pi_1(s,t)$ and $p(s,t)$ can be written in terms of transition probabilities, in the same multi-state model of Figure 3.1:

$$\pi_1(s,t) = P(X_j(s) = 1 \mid T_j > t) = \frac{P_{11}(s,t)P_{01}(0,s)}{P_{00}(0,t) + P_{01}(0,t)}, \tag{3.2.6}$$

$$p(s,t) = P(X_i(s) = 1 \mid T_i > s, T_i \leq t) = \frac{P_{12}(s,t)P_{01}(0,s)}{P_{02}(s,t)P_{00}(0,s) + P_{12}(s,t)P_{01}(0,s)}, \tag{3.2.7}$$

where $P_{kl}(u,v)$ is the conditional probability of being in state l at time v given in state k at time u.

The cumulative/dynamic AUC at time s measures how well the disease marker evaluated at time s separates those who die before time t from those who survive until t.

§3.2.3 AUC for Weibull illness-death model

In this article an illness-death model with Weibull distributed transition hazards is studied because of its simple transition probabilities. The transition hazards from state i to state j are defined by

$$\lambda_{ij}(t) = \alpha_{ij} k t^{k-1}, \tag{3.2.8}$$

where k is the common shape parameter and α_{ij} are transition-specific rate parameters. Let

$$\begin{aligned} S_0(t) &= \exp(-(\alpha_{01} + \alpha_{02})t^k), \\ S_1(t) &= \exp(-\alpha_{12} t^k). \end{aligned}$$

The transition probabilities are then equal to[119]

$$P_{00}(u,t) = \frac{S_0(t)}{S_0(u)},$$

$$P_{11}(u,t) = \frac{S_1(t)}{S_1(u)},$$

$$P_{01}(u,t) = \begin{cases} \dfrac{\alpha_{01}}{\alpha_{01} + \alpha_{02} - \alpha_{12}} \left(\dfrac{S_1(t)}{S_1(u)} - \dfrac{S_0(t)}{S_0(u)} \right) & \text{, if } \alpha_{01} + \alpha_{02} - \alpha_{12} \neq 0 \\ \alpha_{01} \left(\dfrac{S_1(t)}{S_1(u)} t^k - \dfrac{S_0(t)}{S_0(u)} u^k \right) & \text{, otherwise (note: } S_1(t) = S_0(t)), \end{cases}$$

$$P_{02}^0(u,t) = \frac{\alpha_{02}}{\alpha_{01} + \alpha_{02}} \left(1 - \frac{S_0(t)}{S_0(u)} \right),$$

$$P_{02}^1(u,t) = \begin{cases} \frac{\alpha_{01}}{\alpha_{01}+\alpha_{02}}\left(1-\frac{S_0(t)}{S_0(u)}\right) - \frac{\alpha_{01}}{\alpha_{01}+\alpha_{02}-\alpha_{12}}\left(\frac{S_1(t)}{S_1(u)}-\frac{S_0(t)}{S_0(u)}\right), \\ \qquad\qquad \text{if } \alpha_{01}+\alpha_{02}-\alpha_{12} \neq 0 \quad , \\ \frac{\alpha_{01}}{\alpha_{01}+\alpha_{02}}\left(1-\frac{S_0(t)}{S_0(u)}\right) - \alpha_{01}\frac{S_0(t)}{S_0(u)}\left(t^k-u^k\right), \qquad \text{otherwise} \end{cases}$$

$$P_{02}(u,t) = P_{02}^0(u,t) + P_{02}^1(u,t) = 1 - \frac{\alpha_{02}-\alpha_{12}}{\alpha_{01}+\alpha_{02}-\alpha_{12}}\frac{S_0(t)}{S_0(u)} - \frac{\alpha_{01}}{\alpha_{01}+\alpha_{02}-\alpha_{12}}\frac{S_1(t)}{S_1(u)},$$

$$P_{12}(u,t) = 1 - \frac{S_1(t)}{S_1(u)}.$$

These transition probabilities can be used to calculate the time-specific incident/dynamic and cumulative/dynamic AUC using Equations (3.2.1) and (3.2.5), respectively.

§3.2.4 Estimation

Equations (3.2.1)–(3.2.3)and (3.2.5)–(3.2.7) relate, respectively, the incident/dynamic and cumulative/dynamic AUC to transition probabilities and hazards. Estimates for the AUCs can be obtained by replacing transition probabilities and hazards by their estimated counterparts. Such estimates may be obtained from software packages for multi-state models, such as the R-packages `mstate`, `msm`, and `SmoothHazard` discussed in Section 3.3.[49, 48, 85, 144]

§3.2.5 Estimation of incident/dynamic AUC

Equations (3.2.1)–(3.2.3) are used to estimate the incident/dynamic AUC,

$$\widehat{\text{AUC}}^{I/D}(t) = 0.5 + 0.5(\hat{p}(t) - \hat{\pi}_1(t)), \tag{3.2.9}$$

where

$$\hat{\pi}_1(t) = \frac{\hat{P}_{01}(t)}{\hat{P}_{00}(t) + \hat{P}_{01}(t)},$$

$$\hat{p}(t) = \frac{\frac{\hat{\lambda}_{12}(t)}{\hat{\lambda}_{02}(t)}\hat{P}_{01}(t-)}{\hat{P}_{00}(t-) + \frac{\hat{\lambda}_{12}(t)}{\hat{\lambda}_{02}(t)}\hat{P}_{01}(t-)},$$

where $t-$ means just before time t, $\hat{P}_{0l}(t)$ is an estimate of the conditional probability of being in state l, $(l = 0, 1)$ at time t given in state 0 at time 0 and $\hat{\lambda}_{k2}(t)$ is an estimate of the transition hazard at time t for moving from state k, $(k = 0, 1)$ to state 2.

§3.2.6 Estimation of cumulative/dynamic AUC

Equations (3.2.5)–(3.2.7)are used to estimate the cumulative/dynamic AUC,

$$\widehat{\text{AUC}}^{C/D}(s,t) = 0.5 + 0.5(\hat{p}(s,t) - \hat{\pi}_1(s,t)),$$

where

$$\hat{\pi}_1(s,t) = \frac{\hat{P}_{11}(s,t)\hat{P}_{01}(0,s)}{\hat{P}_{00}(0,t) + \hat{P}_{01}(0,t)},$$

$$\hat{p}(s,t) = \frac{\hat{P}_{12}(s,t)\hat{P}_{01}(0,s)}{\hat{P}_{02}(s,t)\hat{P}_{00}(0,s) + \hat{P}_{12}(s,t)\hat{P}_{01}(0,s)},$$

where $\hat{P}_{kl}(u,v)$ is an estimate of the conditional probability of being in state l at time v given in state k at time u.

§3.3 Illness-death models

Four different methods to estimate the illness-death model for interval-censored data were compared: (1) the Cox model with disease state as time-dependent covariate (ignoring the interval-censored nature of the time-dependent covariate), (2) the piecewise-constant model accounting for interval-censoring using the `msm` function from the `msm` package, (3) the Weibull model accounting for interval-censoring using the `idm` function from the `SmoothHazard` package, and (4) the M-spline model accounting for interval-censoring using the `idm` function from the `SmoothHazard` package.[85, 144] A sieve estimator for a Cox based multi-state model that accounts for interval-censoring is implemented in the `coxdual` function from the `coxinterval` package, however, at least some disease times need to be observed exactly for the estimation procedure to work.[31] Since this is not the case in the motivation for this study the `coxinterval` package was not further considered. In the simulation study presented in Section 3.4 all methods are used and from their transition probabilities the AUC is estimated.

§3.3.1 Cox model with time-dependent covariate

The Cox model with a binary time-dependent covariate is defined by the following hazard function:

$$\lambda(t|X(t)) = \lambda_0(t)\exp(\beta X(t)),$$

where $\lambda_0(t)$ is the baseline hazard, $X(t)$ is the binary disease marker at time t and β its effect. This model can be estimated by e.g. the `coxph` R-function from the `survival` package[141], however, ignoring the interval-censored nature of the time-dependent covariate. Disease time is assumed to be the time of diagnosis of disease: $X(t) = 0$ if a patient was not diagnosed with disease yet at time t and $X(t) = 1$ if a patient was diagnosed with disease by time t. The Cox model with time-dependent covariate corresponds to an illness-death model in which the transition hazards to the state death are proportional. This allows for the estimation of the effect of disease on death in form of a hazard ratio (HR). Transition probabilities can be retrieved from the model using `msfit` and `probtrans` functions from the `mstate` package.[49, 48] The `risksetAUC` R-function from the `risksetROC` package[80] estimates the incident/dynamic AUC for a Cox model with time-dependent covariate. Additionally to estimating the AUC using transition probabilities this function is also used in the simulation study in Section 3.4.

§3.3.2 Piecewise-constant model accounting for interval-censoring

This Markov model is described in Figure 3.1. Interval-censored data from an illness-death process are a special case of panel data, in which the state of an individual is observed at a finite series of times. The likelihood for panel data can be calculated in closed form if the transition hazards are constant or piece-wise constant.[85] A model with piecewise-constant hazards given by

$$\lambda_{ij}(t) = \begin{cases} \lambda_{ij1} & \text{, if } t \leq c_1 \\ \lambda_{ij2} & \text{, if } c_1 < t \leq c_2 \\ \vdots \end{cases},$$

where c_k are the times at which the hazard may change is considered. This model is implemented in the `msm` package and can account for the interval-censored disease state.[85] In the simulation study of Section 3.4 the hazards towards the death state are assumed to be proportional so that an effect of disease on survival can be estimated.

§3.3.3 Weibull model accounting for interval-censoring

This model is a Markov illness-death model (see Figure 3.1) which assumes a Weibull distribution for the transition hazards given by

$$\lambda_{ij}(t) = \alpha_{ij} k_{ij} t^{k_{ij}-1},$$

where α_{ij} and k_{ij} are rate and shape parameters for the transition from state i to state j, respectively. This model is implemented in the `SmoothHazard` R-package.[144]

It accounts for interval-censoring and the probability of developing disease between last disease scan and death or lost to follow-up and it is estimated by maximizing the likelihood with the `idm` function. The function does not allow for the transition hazards to the death state to be set proportional and therefore no effect of disease on death can be estimated. The package provides prediction of transition probabilities based on estimated transition hazards.

§3.3.4 M-spline model accounting for interval-censoring

This Markov illness-death model is described in Figure 3.1. The model is estimated using a penalized likelihood approach with non-parametric transition hazards $\lambda_{01}(t), \lambda_{02}(t)$, and $\lambda_{12}(t)$, approximated by M-splines and it is implemented in the `SmoothHazard` R-package.[86, 144] This model as the previous two, accounts for interval-censoring of the disease state as well as the probability of developing disease between the last disease scan and death or lost to follow-up. It is estimated by the `idm` function from the `SmoothHazard` R-package in which the option `method = "Splines"` is set.[144] By default 7 knots per transition are estimated. As for the Weibull model, the transition hazards towards the death state can not be set proportional and therefore no HR for disease can be estimated. Transition probabilities can be obtained using functions provided in the package.

§3.4 Simulation

To study the predictive accuracy of an interval-censored disease marker on survival a simulation study was conducted. Incident/dynamic and cumulative/dynamic AUC were computed to quantify the predictive accuracy of the disease marker for different estimation procedures of the illness-death model. The methods compared were the Cox model with time-dependent disease marker, which ignores interval-censoring, and the illness-death model for interval-censored data estimated with three different implementations: the piecewise-constant model implemented in the `msm` package, the Weibull model, and the M-spline model which are both implemented in the `SmoothHazard` package (see Section 3.3). The piecewise-constant model needs as input pre-specified change points at which the hazard may change. For the simulation study 4 change points were considered 6, 30, 60, and 90 months. For the M-spline model the default of 7 knots per transition was used.

Motivated by the clinical data discussed in Section 3.5 multiple data scenarios were simulated and results from the different methods were compared. The number of individuals per data set was either equal to 1000 or equal to 2000. Data were generated from Weibull transition hazards with a common shape parameter k and different rate parameters α_{01}, α_{02} and α_{12} (see Equation (3.2.8)). The Weibull parameters were based on the data discussed in Section 3.5 and were fixed throughout the simulated scenarios ($\alpha_{01} = 0.05, \alpha_{02} = 0.05, \alpha_{12} = 0.56, k = 0.5$).

The survival time was censored according to two different censoring schemes: either it was censored administratively at 10 years follow-up or censoring times were sampled from a uniform distribution between 5 and 10 years. The disease state was

Table 3.1: Simulated scenarios.

Scenario	N	Censoring	Follow-up
A	1000	unif(5, 10)	3
B	1000	unif(5, 10)	6
C	1000	unif(5, 10)	12
D	1000	10	3
E	1000	10	6
F	1000	10	12
G	2000	unif(5, 10)	3
H	2000	unif(5, 10)	6
I	2000	unif(5, 10)	12
J	2000	10	3
K	2000	10	6
L	2000	10	12

Abbreviations: N, total number of patients; Censoring, type of death censoring, unif(5, 10) means censoring was uniformly sampled between 5 and 10 years and 10 means that administrative censoring occurred at 10 years; Follow-up, time between disease observations in months.

observed only at pre-specified follow-up visits. The scenarios cover three different follow-up schemes in which the disease state was observed every 3, 6, or 12 months. Table 3.1 summarizes the simulated scenarios. Each scenario was simulated 1000 times.

Table 3.2 shows the estimated coefficients and hazard ratios of disease for the piecewise-constant and the Cox model. For the Weibull and M-spline model no effect could be estimated, since the `idm` function does not allow transition hazards to be proportional. The coefficients from the Cox model were consistently more biased than from the piecewise-constant model. The Cox model underestimated the true coefficient and the bias increased for larger follow-up intervals. These results are in line with Leffondré et al. (2013)[97] who showed that the effect estimates of the Cox model were biased if the covariate affected both the risk of disease and death.

Simulation results show that the coefficients from the piecewise-constant model had smaller bias and smaller root mean square error.

AUC results obtained from different methods for scenarios A-F are summarized in Tables 3.3 and 3.4. For results concerning other scenarios, see Appendix 3.B. The cumulative/dynamic AUC was estimated every month and the incident/dynamic AUC was estimated at each event time, because it depends on the transition hazard evaluated at that time. In Table 3.3 where the AUC at specifc time points was investigated, the AUC estimate just before that time was considered.

The M-spline model did not converge for many data sets. In some of these cases this prevented the estimation of the incident/dynamic and cumulative/dynamic AUC.

The number of invalid estimations is shown in Appendix 3.B, Table 3.B.4. The results of the M-spline model in Tables 3.3 and 3.4 are based only on valid estimations. Additionally, for the M-spline model it is not possible to obtain transition probabilities for a time after the last observation time. This restricts the estimation of the cumulative/dynamic AUC (with prediction window of 5 years) to be estimated only until 5 years prior to the last observation time, see Figure 3.2.

Table 3.3 shows the bias, empirical standard error, and root mean square error for estimates of the incident/dynamic AUC at different years. The Weibull model outperformed the other models in every scenario. This is not surprising since data were generated according to Weibull distributions. The M-spline model consistently had the largest standard error as well as the second smallest bias overall. The piecewise-constant model was slightly less biased than the Cox model for scenarios with 6 and 12 months in between follow-up visits (scenarios B, C, E, F). For the scenarios with 3 months in between follow-up visits the Cox model outperformed the piecewise-constant model (scenarios A, D) in terms of bias. The incident/dynamic AUC for the Cox model was estimated by two different approaches. The first approach computes the AUC from the ROC curve derived from the estimated sensitivity and specificity and is implemented in the `risksetAUC` function from the `risksetROC` R-package[80]. The second approach computes the AUC from estimated transition probabilities as described in Equation (3.2.9). Since the two estimation procedures for the Cox model's AUC gave similar result, only results for the transition probability based AUC are presented in Table 3.3 (see Appendix 3.B, Table 3.B.2 for all results).

Table 3.4 shows the bias, empirical standard error, and root mean square error for estimates of the cumulative/dynamic AUC. The piecewise-constant model showed the worst performance and underestimated the true AUC. The Weibull model, M-spline and the Cox model provided good results.

In Table 3.3 and 3.4 the AUC estimates were investigated at 1, 3, and 5 years which coincide with the times of follow-up visits for every scenario. At these times the Cox model displays less bias compared to times in between follow-up visits (see, Figure 3.1 and 3.2).

The censoring scheme did not have a large effect on the incident/dynamic and cumulative/dynamic AUC estimates for the Cox, piecewise-constant and Weibull model. It did however, have an effect on the estimates of the M-spline model. Earlier censoring according to the uniform distribution between 5 and 10 years (scenarios A-C, G-I) resulted in a larger percentage of invalid estimations (see Appendix 3.B, Table 3.B.4), compared to administrative censoring at 10 years (scenarios D-F, J-L).

The number of individuals per data set did not have a large effect on the mean HRs for disease, it did however reduce the empirical standard error (Table 3.2). Average AUC estimates were nearly identical between scenarios where only the size differed and therefore only results for $n = 1000$ are shown in Table 3.3 and 3.4 (see Appendix 3.B for results of all scenarios). The number of patients per data set did have an effect on the percentage of converged M-spline models (see Appendix 3.B, Table 3.B.4).

The follow-up schemes with larger intervals resulted in larger bias of the incident/dynamic AUC estimates, particularly for the Cox model. The follow-up scheme with larger intervals resulted in consistently more biased estimates of the cumulat-

ive/dynamic AUC for the piecewise-constant model. The Cox, Weibull and M-spline model based estimates were of limited bias for the different follow-up schemes.

Figure 3.1 and 3.2 show incident/dynamic and cumulative/dynamic AUC estimates respectively for data scenarios A, B and C with follow-up visits every 3, 6, and 12 months, respectively. Each plot depicts the true AUC in blue and 1000 green lines which correspond to the AUC estimates of each simulated data set. The Cox model's AUC displays jumps at the observation time points. The reason is that at those time points the proportion of diseased individuals is increased in the risk set. Before the first observation time point the curve is equal to 0.5, because no disease was observed yet.

For the incident/dynamic AUC in Figure 3.1 the M-spline model shows a similar behaviour to the Cox model. No distinct jumps are observed but waves can be seen that are most defined at the beginning of follow-up time. Since the piecewise-constant model and the Weibull model make assumptions about the hazard function, the AUC estimates do not display jumps or waves, like for the Cox and M-spline model.

The variation between curves is much larger for the incident/dynamic AUC estimates in Figure 3.1 compared to the cumulative/dynamic estimates in Figure 3.2. Results indicate that the piecewise-constant model is not flexible enough to follow the shape of the true AUC curve, particularly in the incident/dynamic case. The M-spline model displays a larger variance in the incident/dynamic case and shows a better performance in the cumulative/dynamic case. Its cumulative/dynamic curves underestimated the true AUC initially but recovered later on. The Weibull model outperformed the other models, but again one should keep in mind that data were generated from Weibull distributions.

Table 3.2: Effect of disease.

Scenario	N	Censoring	Follow-up	Model	Mean(coef)	exp(mean(coef))	SE(coef)	Bias(coef)	RMSE(coef)
Truth					2.42	11.20			
A	1000	unif(5, 10)	3	Cox	2.35	10.44	0.09	-0.07	0.11
A	1000	unif(5, 10)	3	Piecewise-constant	2.43	11.32	0.09	0.01	0.09
B	1000	unif(5, 10)	6	Cox	2.29	9.91	0.10	-0.12	0.16
B	1000	unif(5, 10)	6	Piecewise-constant	2.48	11.90	0.10	0.06	0.12
C	1000	unif(5, 10)	12	Cox	2.24	9.37	0.11	-0.18	0.21
C	1000	unif(5, 10)	12	Piecewise-constant	2.41	11.11	0.12	-0.01	0.12
D	1000	10	3	Cox	2.35	10.45	0.09	-0.07	0.11
D	1000	10	3	Piecewise-constant	2.45	11.59	0.08	0.03	0.09
E	1000	10	6	Cox	2.31	10.05	0.09	-0.11	0.14
E	1000	10	6	Piecewise-constant	2.50	12.19	0.10	0.08	0.13
F	1000	10	12	Cox	2.25	9.47	0.10	-0.17	0.20
F	1000	10	12	Piecewise-constant	2.43	11.34	0.11	0.01	0.11
G	2000	unif(5, 10)	3	Cox	2.34	10.43	0.07	-0.07	0.10
G	2000	unif(5, 10)	3	Piecewise-constant	2.42	11.30	0.06	0.01	0.06
H	2000	unif(5, 10)	6	Cox	2.29	9.91	0.07	-0.12	0.14
H	2000	unif(5, 10)	6	Piecewise-constant	2.47	11.88	0.07	0.06	0.09
I	2000	unif(5, 10)	12	Cox	2.24	9.38	0.08	-0.18	0.20
I	2000	unif(5, 10)	12	Piecewise-constant	2.41	11.09	0.08	-0.01	0.08
J	2000	10	3	Cox	2.34	10.40	0.06	-0.07	0.10
J	2000	10	3	Piecewise-constant	2.44	11.52	0.06	0.03	0.07
K	2000	10	6	Cox	2.30	10.00	0.06	-0.11	0.13
K	2000	10	6	Piecewise-constant	2.49	12.10	0.07	0.08	0.10
L	2000	10	12	Cox	2.24	9.40	0.07	-0.17	0.19
L	2000	10	12	Piecewise-constant	2.42	11.24	0.08	0.00	0.08

Abbreviations: SE, empirical standard error; RMSE, root mean square error.

Table 3.3: Time-specific incident/dynamic AUC.

		$AUC^{I/D}(1) = 0.71$			$AUC^{I/D}(3) = 0.72$			$AUC^{I/D}(5) = 0.72$		
Scenario	Model	Bias	SE	RMSE	Bias	SE	RMSE	Bias	SE	RMSE
A	Cox	-0.04	0.01	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
A	Piecewise-constant	-0.05	0.01	0.05	-0.03	0.01	0.03	-0.02	0.01	0.03
A	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
A	M-spline	0.02	0.02	0.03	0.00	0.02	0.02	-0.01	0.04	0.04
B	Cox	-0.07	0.02	0.08	-0.04	0.01	0.04	-0.03	0.01	0.03
B	Piecewise-constant	-0.06	0.01	0.07	-0.04	0.01	0.04	-0.03	0.01	0.03
B	Weibull	-0.01	0.02	0.02	0.00	0.01	0.01	0.00	0.01	0.01
B	M-spline	0.01	0.03	0.03	-0.01	0.03	0.03	0.00	0.04	0.04
C	Cox	-0.21	0.00	0.21	-0.07	0.01	0.07	-0.06	0.01	0.06
C	Piecewise-constant	-0.09	0.01	0.09	-0.06	0.01	0.06	-0.05	0.02	0.05
C	Weibull	-0.01	0.03	0.03	0.00	0.02	0.02	0.00	0.02	0.02
C	M-spline	-0.05	0.04	0.06	0.00	0.03	0.03	0.00	0.04	0.04
D	Cox	-0.04	0.01	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
D	Piecewise-constant	-0.05	0.01	0.05	-0.03	0.01	0.03	-0.02	0.01	0.03
D	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
D	M-spline	0.02	0.02	0.03	-0.01	0.02	0.03	0.00	0.03	0.04
E	Cox	-0.07	0.02	0.07	-0.04	0.01	0.04	-0.03	0.01	0.03
E	Piecewise-constant	-0.06	0.01	0.06	-0.04	0.01	0.04	-0.03	0.01	0.03
E	Weibull	-0.01	0.02	0.02	0.00	0.01	0.01	0.00	0.01	0.01
E	M-spline	0.01	0.02	0.03	-0.01	0.03	0.03	0.00	0.04	0.04
F	Cox	-0.21	0.00	0.21	-0.07	0.01	0.07	-0.06	0.01	0.06
F	Piecewise-constant	-0.09	0.01	0.09	-0.06	0.01	0.06	-0.05	0.01	0.05
F	Weibull	-0.01	0.02	0.02	0.00	0.01	0.01	0.00	0.01	0.01
F	M-spline	-0.05	0.04	0.06	0.00	0.03	0.03	0.00	0.04	0.04

Abbreviations: $AUC^{I/D}(x)$, incident/dynamic AUC at year x; SE, empirical standard error; RMSE, root mean square error.

Table 3.4: Time-specific cumulative/dynamic AUC.

		$AUC^{C/D}(1) = 0.59$			$AUC^{C/D}(3) = 0.62$			$AUC^{C/D}(5) = 0.64$		
Scenario	Model	Bias	SE	RMSE	Bias	SE	RMSE	Bias	SE	RMSE
A	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
A	Piecewise-constant	-0.02	0.01	0.02	-0.01	0.01	0.02	-0.01	0.01	0.02
A	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
A	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
B	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
B	Piecewise-constant	-0.04	0.00	0.04	-0.02	0.01	0.03	-0.02	0.01	0.02
B	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
B	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
C	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
C	Piecewise-constant	-0.05	0.00	0.05	-0.04	0.01	0.04	-0.03	0.01	0.03
C	Weibull	-0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
C	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
D	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
D	Piecewise-constant	-0.02	0.01	0.02	-0.01	0.01	0.02	-0.01	0.01	0.02
D	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
D	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
E	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
E	Piecewise-constant	-0.04	0.00	0.04	-0.03	0.01	0.03	-0.02	0.01	0.02
E	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
E	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
F	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
F	Piecewise-constant	-0.05	0.00	0.05	-0.04	0.01	0.04	-0.03	0.01	0.03
F	Weibull	-0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
F	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02

Abbreviations: $AUC^{C/D}(x)$, cumulative/dynamic AUC at year x; SE, empirical standard error; RMSE, root mean square error.

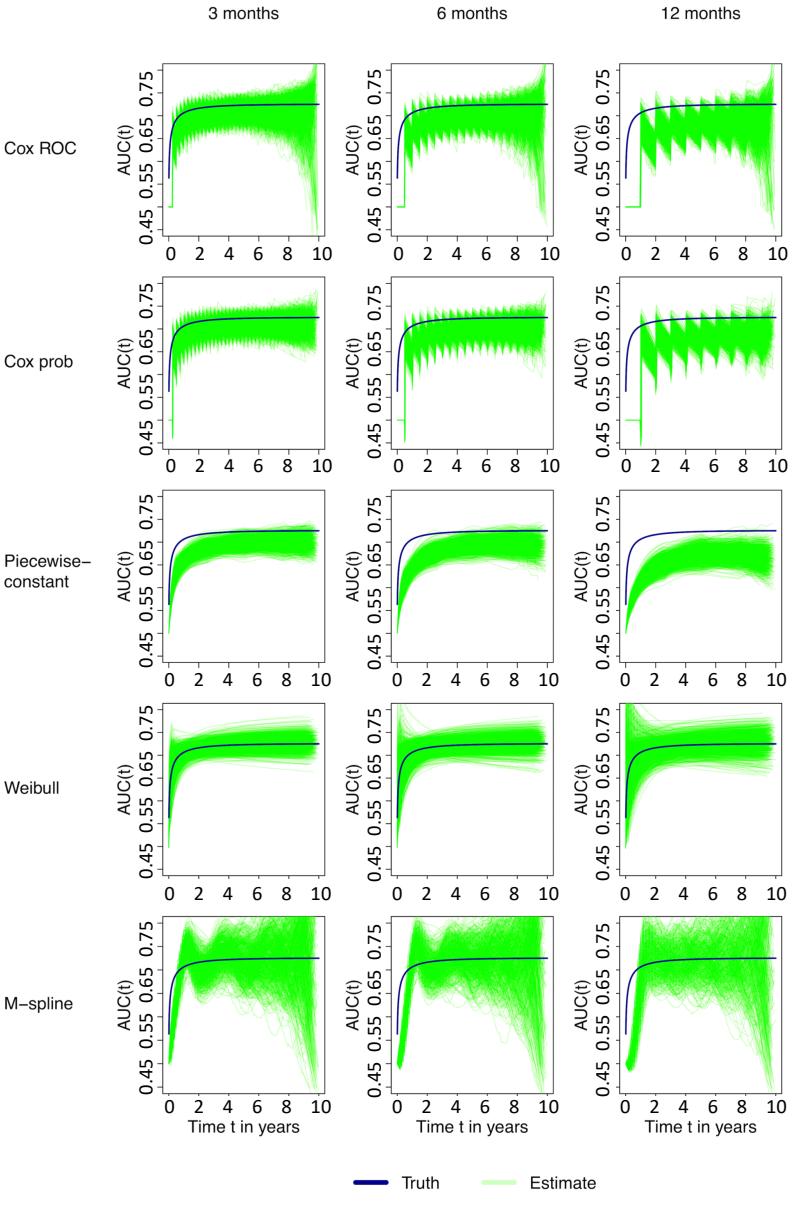

Figure 3.1: Incident/dynamic AUC for scenario A (3 months), B (6 months) and C (12 months). ***Abbreviations:*** *Cox ROC, estimate based on* `risksetAUC` *function; Cox prob, estimate based on transition probabilities of Cox model.*

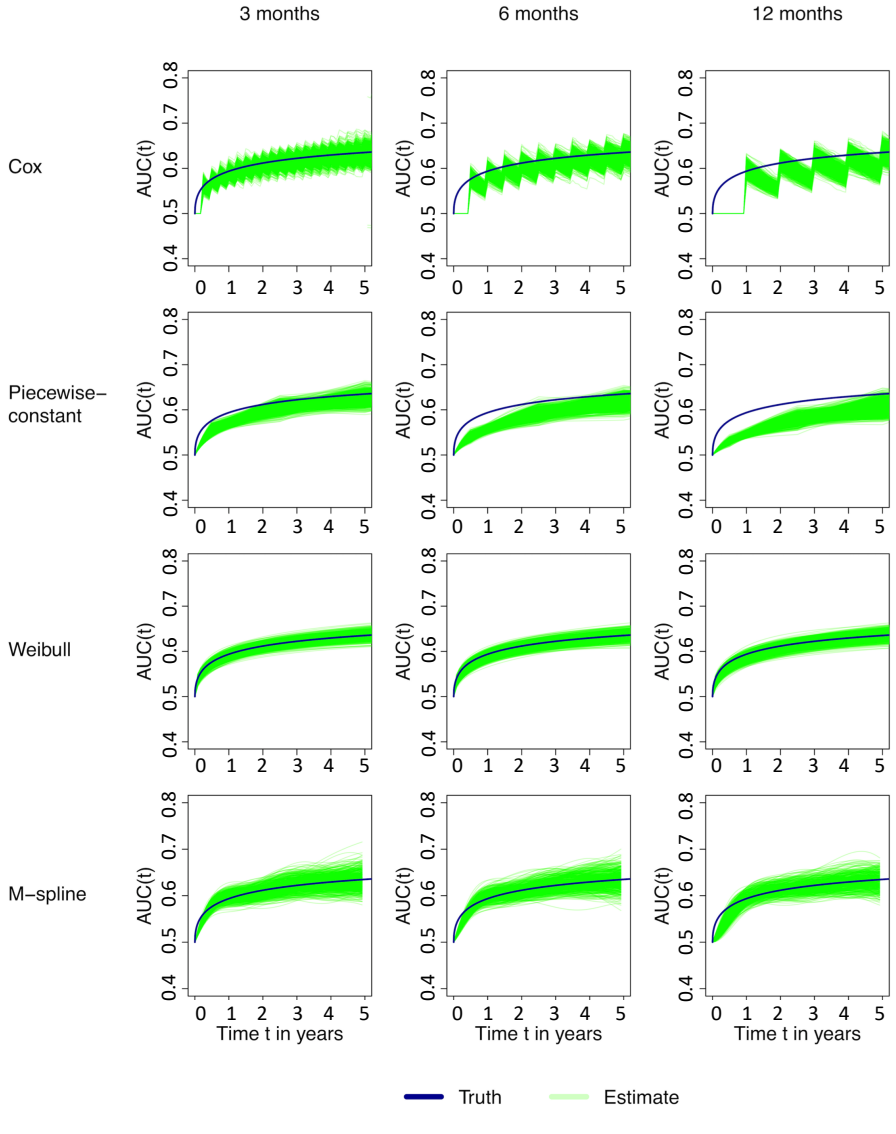

Figure 3.2: Cumulative/dynamic AUC for scenario A (3 months), B (6 months) and C (12 months).

§3.5 Application

The data analyzed in this section was used for the development of a dynamic prediction model for high-grade soft tissue sarcoma patients.[19] The data set contains follow-up information of 2232 patients treated surgically with curative intent. Median follow-up time was 6.42 years. After surgery disease progression can be described by several adverse events: a patient may develop a local recurrence and/or develop distant metastasis (DM) and/or die. The analysis discussed in this section focuses on the effect of DM on death. In total 1034 patients died and 715 patients developed DM (see Figure 3.1).

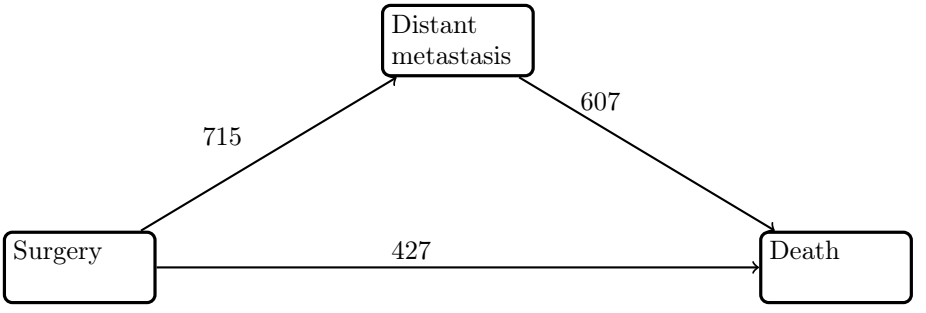

Figure 3.1: Soft tissue sarcoma illness-death model (n = 2232).

After surgery a common follow-up visit scheme to screen for DM is to see a patient every 3 months within the first 3 years, then every 6 months until year 5, and from then on once a year.[66] The data did not contain information about exact follow-up times and an approximation of disease screening times was applied. For a patient who was diagnosed with DM during follow-up, the time of DM was interpreted as the first positive screening for DM. Depending on whether DM was diagnosed within the first 3 years, between 3 and 5, or after 5 years the previous screening was assumed to have taken place either 3, 6, or 12 months prior. A patient who was never diagnosed with DM was assumed to have been screened according to the common follow-up scheme described above.

Table 3.1 shows HRs for DM and estimates for the time-specific AUC at different years. The HRs estimated by the Cox and piecewise-constant model are similar, with HRs for the Cox model being slightly larger.

Figure 3.2 displays on the left and the right panel the non-parametric cumulative baseline hazards and a graphical check of their fit to a Weibull distribution, respectively. For this figure the time of DM was assumed to be equal to the time that DM was detected during screening. If the hazards were coming from a Weibull distribution the lines in the right panel of Figure 3.2 would be straight, which is not the case in particular for the transition from surgery to DM. The Weibull model therefore may not be appropriate for this data.

Figure 3.3 shows the AUC over time for the different models. The incident/dynamic AUC of the Weibull model is initially much larger compared to the other models and declines over time. The incident/dynamic AUC of the piecewise-constant model is

the lowest of all three methods. The cumulative/dynamic AUC of the Cox model is generally the largest and the Weibull models the lowest. The M-spline model did not converge for this data set and consequently the incident/dynamic and cumulative/dynamic AUC could not be estimated.

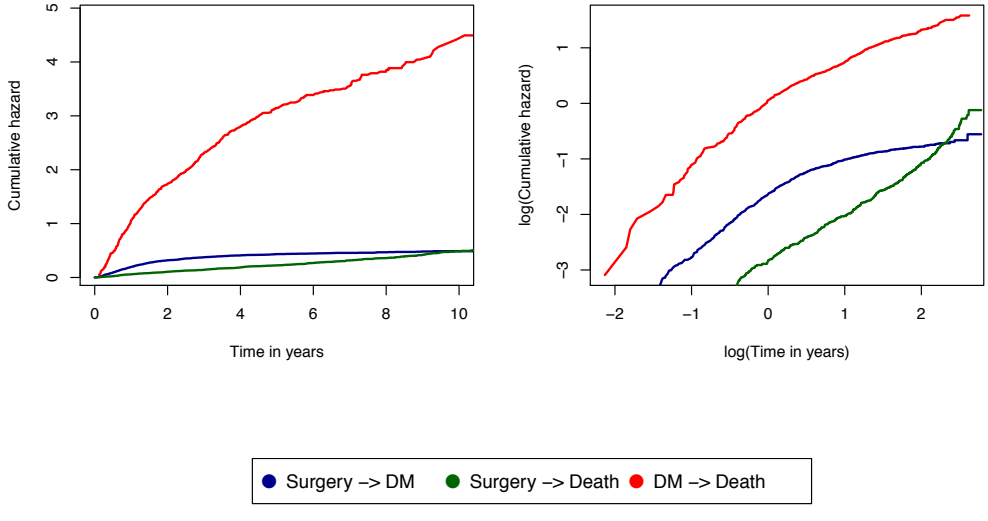

Figure 3.2: Left panel: Cumulative transition hazards. Right panel: plot of ln[H(x)] vs. ln(x) to empirically check the fit of the Weibull distribution.

Table 3.1: Effect and predictive accuracy for distant metastasis.

	HR(DM)	$AUC^{I/D}(1)$	$AUC^{I/D}(2)$	$AUC^{I/D}(3)$	$AUC^{I/D}(4)$	$AUC^{I/D}(5)$
Cox ROC	11.71	0.74	0.76	0.76	0.75	0.74
Cox prob	11.71	0.75	0.76	0.76	0.75	0.74
Piecewise-constant	11.28	0.71	0.73	0.73	0.71	0.70
Weibull		0.81	0.78	0.76	0.74	0.72
		$AUC^{C/D}(1)$	$AUC^{C/D}(2)$	$AUC^{C/D}(3)$	$AUC^{C/D}(4)$	$AUC^{C/D}(5)$
Cox ROC						
Cox prob		0.64	0.69	0.70	0.68	0.67
Piecewise-constant		0.62	0.66	0.68	0.66	0.66
Weibull		0.62	0.63	0.64	0.64	0.64

Abbreviations: $AUC^{I/D}(x)$, incident/dynamic AUC at year x; $AUC^{C/D}(x)$, cumulative/dynamic AUC at year x; Cox ROC, estimate based on Cox model through `risksetAUC` function; Cox prob, estimate based on Cox model through transition probabilities; HR(DM), hazard ratio of DM.

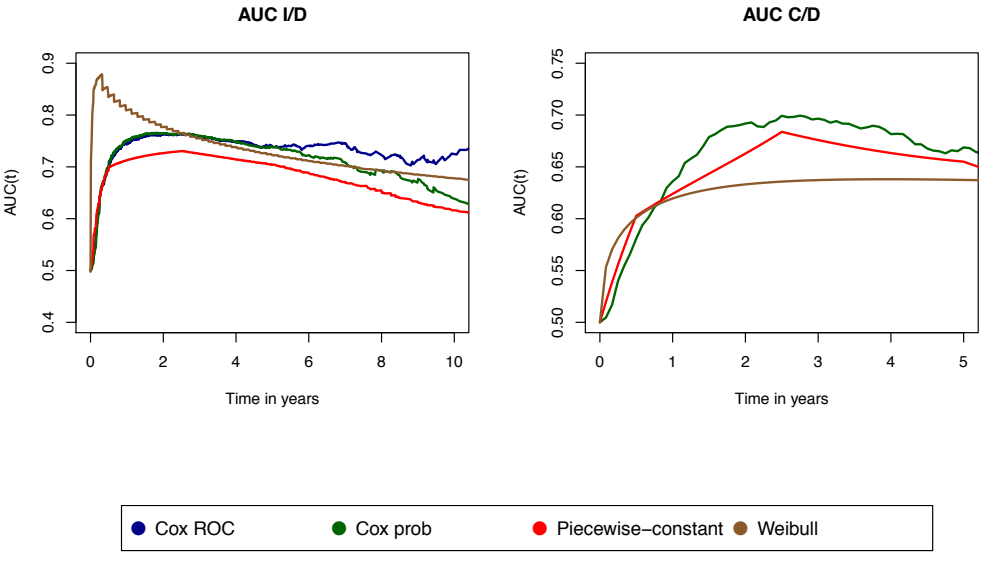

Figure 3.3: Time-specific AUC for distant metastasis. ***Abbreviations:*** *AUC I/D, incident/dynamic AUC; AUC C/D, cumulative/dynamic AUC; Cox ROC, estimate based on Cox model through* `risksetAUC` *function; Cox prob, estimate based on Cox model through transition probabilities.*

§3.6 Discussion

The illness-death model is frequently applied to clinical data to describe disease progression. A patient enters the model disease free, he can then experience disease and die. In clinical practice however, often the time of disease cannot be observed exactly. The information is interval-censored or unobserved because of death or censoring. This can lead to bias in the estimation of disease incidence and regression coefficients.[86, 97]

This article studied the predictive accuracy of a binary time-dependent disease marker in the context of the illness-death model for interval-censored data. A simulation study with several data scenarios was conducted to study four different models: the Cox model with disease as time-dependent marker, the piecewise-constant model implemented in the `msm` package, the Weibull model, and the M-spline model implemented in the `SmoothHazard` package. Both incident/dynamic and cumulative/dynamic AUC estimates were derived from their transition probabilities and studied. The methods were applied to a data set of soft tissue sarcoma patients who were scanned for distant metastasis at scheduled follow-up visits.

The simulation study showed that the HRs from the piecewise-constant model were less biased than those of the Cox model. The number of patients per data set (1000 vs 2000) did not have a large effect on the estimates of the HR, AUC estimates in incident/dynamic and cumulative/dynamic definition except for the M-spline model. The M-spline model converged more reliably with large data sets. The

spacing of follow-up visits at which the disease state was observed did have a large effect on estimates of the incident/dynamic AUC. The Weibull model showed the best performance, however this model had an unfair advantage since the simulated data had Weibull distribution. In practice a Weibull distribution may not be a good fit to the data. The M-spline model showed a good performance when estimating the incident/dynamic and cumulative/dynamic AUC however was not always able to converge and provide AUC estimates. The piecewise-constant model under performed. Even though, incident/dynamic AUC estimates had less bias than the Cox model's for scenarios with large spacing between follow-up visits, cumulative/dynamic estimates had the largest bias of all methods.

Prediction models are becoming more and more important in clinical practice to provide individualized patient care. Dynamic prediction models can incorporate time-dependent disease markers and the predictive accuracy of such a marker may be of interest. In the presence of interval-censored disease time, the results of this study suggest to take the interval-censoring into account not only when estimating parameters of the model, but also when evaluating the predictive accuracy of disease.

Simulations performed studied the effect of an interval-censored binary disease marker. Future research should focus on the predictive accuracy of a time-dependent covariate that can take more than 2 values as well as continuous markers.

Appendix

§3.A Derivation of AUC

§3.A.1 Incident/dynamic AUC

Let i, j be individuals, $X_i(t), X_j(t)$ the binary covariate values at time t, and T_i and T_j the death times. The incident/dynamic AUC is defined as

$$\begin{aligned}
\text{AUC}^{I/D}(t) =& P(X_i(t) > X_j(t) \mid T_i = t, T_j > t) + 0.5P(X_i(t) = X_j(t) \mid T_i = t, T_j > t) \\
=& P(X_j(t) = 0 \mid T_j > t)P(X_i(t) = 1 \mid T_i = t) \\
&+ 0.5[P(X_j(t) = 0 \mid T_j > t)P(X_i(t) = 0 \mid T_i = t) + \\
&\quad P(X_j(t) = 1 \mid T_j > t)P(X_i(t) = 1 \mid T_i = t)] \\
=& (1 - P(X_j(t) = 1 \mid T_j > t))P(X_i(t) = 1 \mid T_i = t) \\
&+ 0.5[(1 - P(X_j(t) = 1 \mid T_j > t))(1 - P(X_i(t) = 1 \mid T_i = t)) + \\
&\quad P(X_j(t) = 1 \mid T_j > t)P(X_i(t) = 1 \mid T_i = t)] \\
=& (1 - \pi_1(t))p(t) + 0.5[(1 - \pi_1(t))(1 - p(t)) + \pi_1(t)p(t)] \\
=& p(t) - \pi_1(t)p(t) + 0.5[1 - p(t) - \pi_1(t) + \pi_1(t)p(t) + \pi_1(t)p(t)] \\
=& p(t) - \pi_1(t)p(t) + 0.5 - 0.5p(t) - 0.5\pi_1(t) + \pi_1(t)p(t) \\
=& 0.5 + 0.5(p(t) - \pi_1(t)),
\end{aligned}$$

where

$$\begin{aligned}
\pi_1(t) =& P(X_i(t) = 1 \mid T_i > t) \\
p(t) =& P(X_i(t-) = 1 \mid T_i = t).
\end{aligned}$$

§3.A.2 Cumulative/dynamic AUC

Let i, j be individuals, $X_i(s), X_j(s)$ their binary covariate values at time s, and T_i and T_j their death times. The cumulative/dynamic AUC is then

$$
\begin{aligned}
\mathrm{AUC}^{C/D}(s,t) =& P(X_i(s) > X_j(s) \mid T_i > s, T_i \le t, T_j > t) + \\
& 0.5P(X_i(s) = X_j(s) \mid T_i > s, T_i \le t, T_j > t) \\
=& P(X_j(s) = 0 \mid T_j > t)P(X_i(s) = 1 \mid T_i > s, T_i \le t) \\
& + 0.5[P(X_j(s) = 0 \mid T_j > t)P(X_i(s) = 0 \mid T_i > s, T_i \le t) + \\
& P(X_j(s) = 1 \mid T_j > t)P(X_i(s) = 1 \mid T_i > s, T_i \le t)] \\
=& (1 - P(X_j(s) = 1 \mid T_j > t))P(X_i(s) = 1 \mid T_i > s, T_i \le t) \\
& + 0.5[(1 - P(X_j(s) = 1 \mid T_j > t))(1 - P(X_i(s) = 1 \mid T_i > s, T_i \le t)) + \\
& P(X_j(s) = 1 \mid T_j > t)P(X_i(s) = 1 \mid T_i > s, T_i \le t)] \\
=& (1 - \pi_1(s,t))p(s,t) + 0.5[(1 - \pi_1(s,t))(1 - p(s,t)) + \pi_1(s,t)p(s,t)] \\
=& p(s,t) - \pi_1(s,t)p(s,t) + 0.5[1 - p(s,t) - \pi_1(s,t) + \pi_1(s,t)p(s,t) + \pi_1(s,t)p(s,t)] \\
=& p(s,t) - \pi_1(s,t)p(s,t) + 0.5 - 0.5p(s,t) - 0.5\pi_1(s,t) + \pi_1(s,t)p(s,t) \\
=& 0.5 + 0.5(p(s,t) - \pi_1(s,t)),
\end{aligned}
$$

where

$$
\pi_1(s,t) = P(X_j(s) = 1 \mid T_j > t)
$$

$$
p(s,t) = P(X_i(s) = 1 \mid T_i > s, T_i \le t).
$$

§3.B Results for all scenarios

Table 3.B.1: Effect of disease.

Scenario	N	Censoring	Follow-up	Model	Mean(coef)	exp(mean(coef))	SE(coef)	Bias(coef)	RMSE(coef)
Truth					2.42	11.20			
A	1000	unif(5, 10)	3	Cox ROC	2.35	10.44	0.09	-0.07	0.11
A	1000	unif(5, 10)	3	Cox prob	2.35	10.44	0.09	-0.07	0.11
A	1000	unif(5, 10)	3	Piecewise-constant	2.43	11.32	0.09	0.01	0.09
B	1000	unif(5, 10)	6	Cox ROC	2.29	9.91	0.10	-0.12	0.16
B	1000	unif(5, 10)	6	Cox prob	2.29	9.91	0.10	-0.12	0.16
B	1000	unif(5, 10)	6	Piecewise-constant	2.48	11.90	0.10	0.06	0.12
C	1000	unif(5, 10)	12	Cox ROC	2.24	9.37	0.11	-0.18	0.21
C	1000	unif(5, 10)	12	Cox prob	2.24	9.37	0.11	-0.18	0.21
C	1000	unif(5, 10)	12	Piecewise-constant	2.41	11.11	0.12	-0.01	0.12
D	1000	10	3	Cox ROC	2.35	10.45	0.09	-0.07	0.11
D	1000	10	3	Cox prob	2.35	10.45	0.09	-0.07	0.11
D	1000	10	3	Piecewise-constant	2.45	11.59	0.08	0.03	0.09
E	1000	10	6	Cox ROC	2.31	10.05	0.09	-0.11	0.14
E	1000	10	6	Cox prob	2.31	10.05	0.09	-0.11	0.14
E	1000	10	6	Piecewise-constant	2.50	12.19	0.10	0.08	0.13
F	1000	10	12	Cox ROC	2.25	9.47	0.10	-0.17	0.20
F	1000	10	12	Cox prob	2.25	9.47	0.10	-0.17	0.20
F	1000	10	12	Piecewise-constant	2.43	11.34	0.11	0.01	0.11
G	2000	unif(5, 10)	3	Cox ROC	2.34	10.43	0.07	-0.07	0.10
G	2000	unif(5, 10)	3	Cox prob	2.34	10.43	0.07	-0.07	0.10
G	2000	unif(5, 10)	3	Piecewise-constant	2.42	11.30	0.06	0.01	0.06
H	2000	unif(5, 10)	6	Cox ROC	2.29	9.91	0.07	-0.12	0.14
H	2000	unif(5, 10)	6	Cox prob	2.29	9.91	0.07	-0.12	0.14
H	2000	unif(5, 10)	6	Piecewise-constant	2.47	11.88	0.07	0.06	0.09
I	2000	unif(5, 10)	12	Cox ROC	2.24	9.38	0.08	-0.18	0.20
I	2000	unif(5, 10)	12	Cox prob	2.24	9.38	0.08	-0.18	0.20
I	2000	unif(5, 10)	12	Piecewise-constant	2.41	11.09	0.08	-0.01	0.08
J	2000	10	3	Cox ROC	2.34	10.40	0.06	-0.07	0.10
J	2000	10	3	Cox prob	2.34	10.40	0.06	-0.07	0.10
J	2000	10	3	Piecewise-constant	2.44	11.52	0.06	0.03	0.07
K	2000	10	6	Cox ROC	2.30	10.00	0.06	-0.11	0.13
K	2000	10	6	Cox prob	2.30	10.00	0.06	-0.11	0.13
K	2000	10	6	Piecewise-constant	2.49	12.10	0.07	0.08	0.10
L	2000	10	12	Cox ROC	2.24	9.40	0.07	-0.17	0.19
L	2000	10	12	Cox prob	2.24	9.40	0.07	-0.17	0.19
L	2000	10	12	Piecewise-constant	2.42	11.24	0.08	0.00	0.08

Abbreviations: SE, empirical standard error; RMSE, root mean square error.

Table 3.B.2: Time-specific incident/dynamic AUC.

		$\text{AUC}^{I/D}(1) = 0.71$			$\text{AUC}^{I/D}(3) = 0.72$			$\text{AUC}^{I/D}(5) = 0.72$		
Scenario	Model	Bias	SE	RMSE	Bias	SE	RMSE	Bias	SE	RMSE
A	Cox ROC	-0.04	0.02	0.04	-0.02	0.01	0.03	-0.02	0.01	0.02
A	Cox prob	-0.04	0.01	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
A	Piecewise-constant	-0.05	0.01	0.05	-0.03	0.01	0.03	-0.02	0.01	0.03

Table 3.B.2: (continued)

		$AUC^{I/D}(1) = 0.71$			$AUC^{I/D}(3) = 0.72$			$AUC^{I/D}(5) = 0.72$		
Scenario	Model	Bias	SE	RMSE	Bias	SE	RMSE	Bias	SE	RMSE
A	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
A	M-spline	0.02	0.02	0.03	0.00	0.02	0.02	-0.01	0.04	0.04
B	Cox ROC	-0.07	0.02	0.08	-0.04	0.01	0.04	-0.03	0.01	0.04
B	Cox prob	-0.07	0.02	0.08	-0.04	0.01	0.04	-0.03	0.01	0.03
B	Piecewise-constant	-0.06	0.01	0.07	-0.04	0.01	0.04	-0.03	0.01	0.03
B	Weibull	-0.01	0.02	0.02	0.00	0.01	0.01	0.00	0.01	0.01
B	M-spline	0.01	0.03	0.03	-0.01	0.03	0.03	0.00	0.04	0.04
C	Cox ROC	-0.21	0.00	0.21	-0.07	0.02	0.08	-0.06	0.02	0.06
C	Cox prob	-0.21	0.00	0.21	-0.07	0.01	0.07	-0.06	0.01	0.06
C	Piecewise-constant	-0.09	0.01	0.09	-0.06	0.01	0.06	-0.05	0.02	0.05
C	Weibull	-0.01	0.03	0.03	0.00	0.02	0.02	0.00	0.02	0.02
C	M-spline	-0.05	0.04	0.06	0.00	0.03	0.03	0.00	0.04	0.04
D	Cox ROC	-0.04	0.02	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
D	Cox prob	-0.04	0.01	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
D	Piecewise-constant	-0.05	0.01	0.05	-0.03	0.01	0.03	-0.02	0.01	0.03
D	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
D	M-spline	0.02	0.02	0.03	-0.01	0.02	0.03	0.00	0.03	0.04
E	Cox ROC	-0.07	0.02	0.08	-0.04	0.01	0.04	-0.03	0.02	0.03
E	Cox prob	-0.07	0.02	0.07	-0.04	0.01	0.04	-0.03	0.01	0.03
E	Piecewise-constant	-0.06	0.01	0.06	-0.04	0.01	0.04	-0.03	0.01	0.03
E	Weibull	-0.01	0.02	0.02	0.00	0.01	0.01	0.00	0.01	0.01
E	M-spline	0.01	0.02	0.03	-0.01	0.03	0.03	0.00	0.04	0.04
F	Cox ROC	-0.21	0.00	0.21	-0.07	0.02	0.08	-0.06	0.02	0.06
F	Cox prob	-0.21	0.00	0.21	-0.07	0.01	0.07	-0.06	0.01	0.06
F	Piecewise-constant	-0.09	0.01	0.09	-0.06	0.01	0.06	-0.05	0.01	0.05
F	Weibull	-0.01	0.02	0.02	0.00	0.01	0.01	0.00	0.01	0.01
F	M-spline	-0.05	0.04	0.06	0.00	0.03	0.03	0.00	0.04	0.04
G	Cox ROC	-0.04	0.01	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
G	Cox prob	-0.04	0.01	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
G	Piecewise-constant	-0.05	0.01	0.05	-0.03	0.01	0.03	-0.02	0.01	0.03
G	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
G	M-spline	0.02	0.02	0.03	0.00	0.02	0.02	-0.01	0.02	0.03
H	Cox ROC	-0.07	0.01	0.08	-0.04	0.01	0.04	-0.03	0.01	0.03
H	Cox prob	-0.07	0.01	0.08	-0.04	0.01	0.04	-0.03	0.01	0.03
H	Piecewise-constant	-0.06	0.01	0.06	-0.04	0.01	0.04	-0.03	0.01	0.03
H	Weibull	-0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
H	M-spline	0.01	0.02	0.02	-0.01	0.02	0.02	-0.01	0.03	0.03
I	Cox ROC	-0.21	0.00	0.21	-0.07	0.01	0.07	-0.06	0.01	0.06
I	Cox prob	-0.21	0.00	0.21	-0.07	0.01	0.07	-0.06	0.01	0.06
I	Piecewise-constant	-0.09	0.01	0.09	-0.06	0.01	0.06	-0.05	0.01	0.05
I	Weibull	-0.01	0.02	0.02	0.00	0.01	0.01	0.00	0.01	0.01
I	M-spline	-0.04	0.03	0.05	0.00	0.02	0.02	0.00	0.03	0.03
J	Cox ROC	-0.04	0.01	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
J	Cox prob	-0.04	0.01	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
J	Piecewise-constant	-0.05	0.01	0.05	-0.03	0.01	0.03	-0.02	0.01	0.03
J	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
J	M-spline	0.02	0.01	0.03	-0.01	0.02	0.02	-0.01	0.02	0.03
K	Cox ROC	-0.07	0.01	0.07	-0.04	0.01	0.04	-0.03	0.01	0.03
K	Cox prob	-0.07	0.01	0.07	-0.04	0.01	0.04	-0.03	0.01	0.03
K	Piecewise-constant	-0.06	0.01	0.06	-0.04	0.01	0.04	-0.03	0.01	0.03
K	Weibull	-0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
K	M-spline	0.01	0.02	0.02	-0.01	0.02	0.02	0.00	0.02	0.03
L	Cox ROC	-0.21	0.00	0.21	-0.07	0.01	0.07	-0.06	0.01	0.06
L	Cox prob	-0.21	0.00	0.21	-0.07	0.01	0.07	-0.06	0.01	0.06
L	Piecewise-constant	-0.09	0.01	0.09	-0.06	0.01	0.06	-0.05	0.01	0.05
L	Weibull	-0.01	0.02	0.02	0.00	0.01	0.01	0.00	0.01	0.01
L	M-spline	-0.04	0.03	0.05	0.00	0.02	0.02	0.00	0.03	0.03

Abbreviations: $AUC^{I/D}$(x), incident/dynamic AUC at year x; SE, empirical standard error; RMSE, root mean square error.

Table 3.B.3: (continued)

		$AUC^{C/D}(1) = 0.59$			$AUC^{C/D}(3) = 0.62$			$AUC^{C/D}(5) = 0.64$		
Scenario	Model	Bias	SE	RMSE	Bias	SE	RMSE	Bias	SE	RMSE

Table 3.B.3: Time-specific cumulative/dynamic AUC.

		$AUC^{C/D}(1) = 0.59$			$AUC^{C/D}(3) = 0.62$			$AUC^{C/D}(5) = 0.64$		
Scenario	Model	Bias	SE	RMSE	Bias	SE	RMSE	Bias	SE	RMSE
A	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
A	Piecewise-constant	-0.02	0.01	0.02	-0.01	0.01	0.02	-0.01	0.01	0.02
A	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
A	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
B	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
B	Piecewise-constant	-0.04	0.00	0.04	-0.02	0.01	0.03	-0.02	0.01	0.02
B	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
B	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
C	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
C	Piecewise-constant	-0.05	0.00	0.05	-0.04	0.01	0.04	-0.03	0.01	0.03
C	Weibull	-0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
C	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
D	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
D	Piecewise-constant	-0.02	0.01	0.02	-0.01	0.01	0.02	-0.01	0.01	0.02
D	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
D	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
E	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
E	Piecewise-constant	-0.04	0.00	0.04	-0.03	0.01	0.03	-0.02	0.01	0.02
E	Weibull	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
E	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
F	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
F	Piecewise-constant	-0.05	0.00	0.05	-0.04	0.01	0.04	-0.03	0.01	0.03
F	Weibull	-0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
F	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.02	0.02
G	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
G	Piecewise-constant	-0.02	0.00	0.02	-0.01	0.01	0.01	-0.01	0.01	0.01
G	Weibull	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01
G	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
H	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
H	Piecewise-constant	-0.04	0.00	0.04	-0.02	0.01	0.02	-0.02	0.01	0.02
H	Weibull	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.01	0.01
H	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
I	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
I	Piecewise-constant	-0.05	0.00	0.05	-0.04	0.01	0.04	-0.03	0.01	0.03
I	Weibull	-0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
I	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
J	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
J	Piecewise-constant	-0.02	0.00	0.02	-0.01	0.01	0.01	-0.01	0.01	0.01
J	Weibull	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
J	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
K	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
K	Piecewise-constant	-0.04	0.00	0.04	-0.02	0.01	0.03	-0.02	0.01	0.02
K	Weibull	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01
K	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
L	Cox	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
L	Piecewise-constant	-0.05	0.00	0.05	-0.04	0.01	0.04	-0.03	0.01	0.03
L	Weibull	-0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
L	M-spline	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01

Abbreviations: $AUC^{C/D}(x)$, cumulative/dynamic AUC at year x; SE, empirical standard error; RMSE, root mean square error.

Table 3.B.4: Invalid estimations for M-spline model.

	Scenario											
	A	B	C	D	E	F	G	H	I	J	K	L
invalid $AUC^{I/D}$	379	442	555	17	34	88	201	292	378	1	7	19
invalid $AUC^{C/D}$	381	444	560	17	34	89	201	296	381	1	7	19

Number of invalid estimations of $AUC^{I/D}$ and $AUC^{C/D}$ from 1 year based on 1000 data sets.

§3.C Incident/dynamic AUC for scenarios D–L

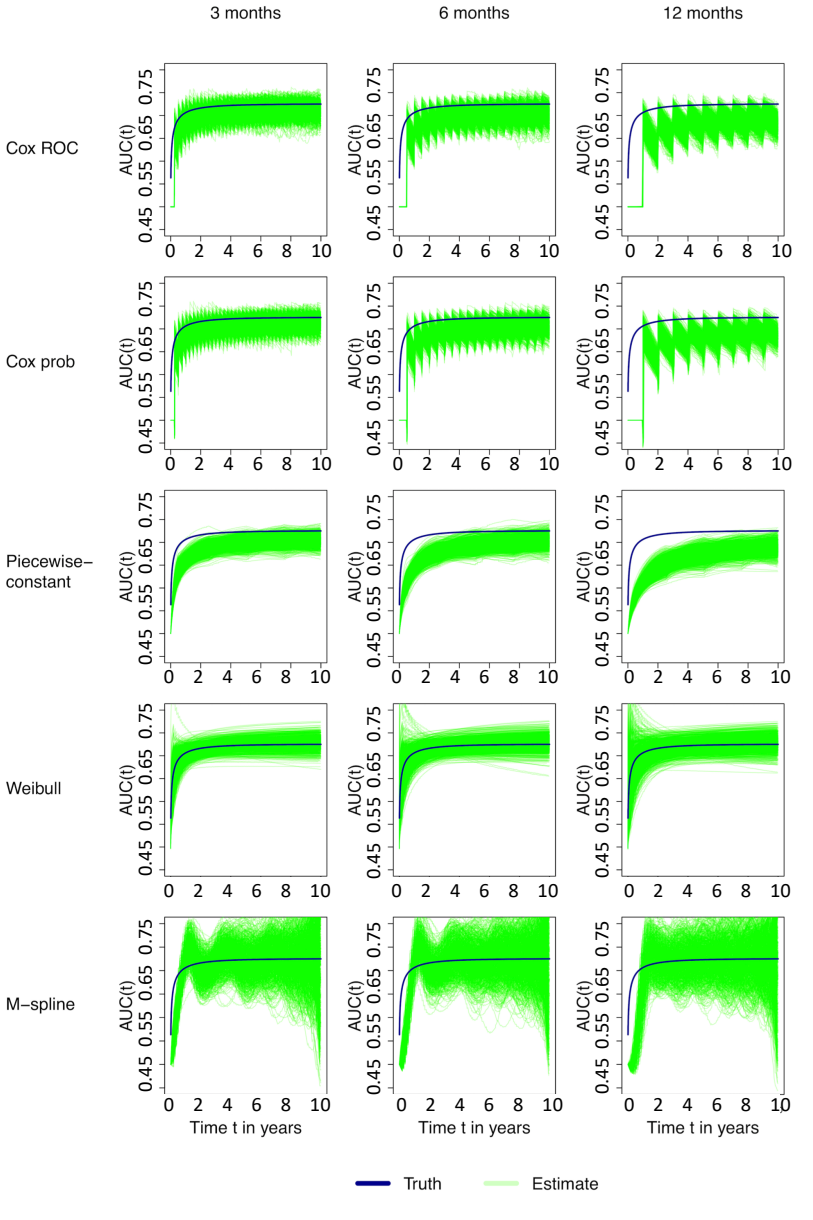

Figure 3.C.1: Incident/dynamic AUC for scenario D (3 months), E (6 months) and F (12 months). ***Abbreviations:*** *Cox ROC, estimate based on* `risksetAUC` *function; Cox prob, estimate based on transition probabilities of Cox model.*

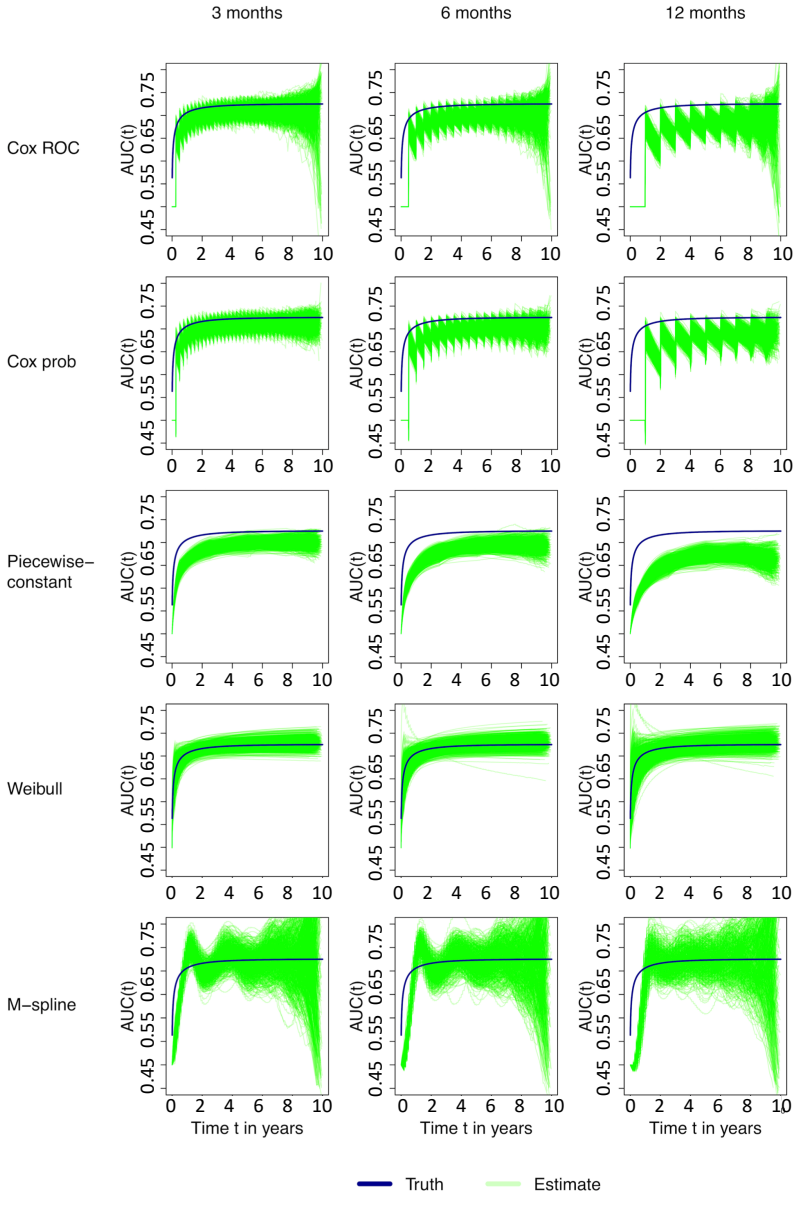

Figure 3.C.2: Incident/dynamic AUC for scenario G (3 months), H (6 months) and I (12 months). ***Abbreviations:*** *Cox ROC, estimate based on* `risksetAUC` *function; Cox prob, estimate based on transition probabilities of Cox model.*

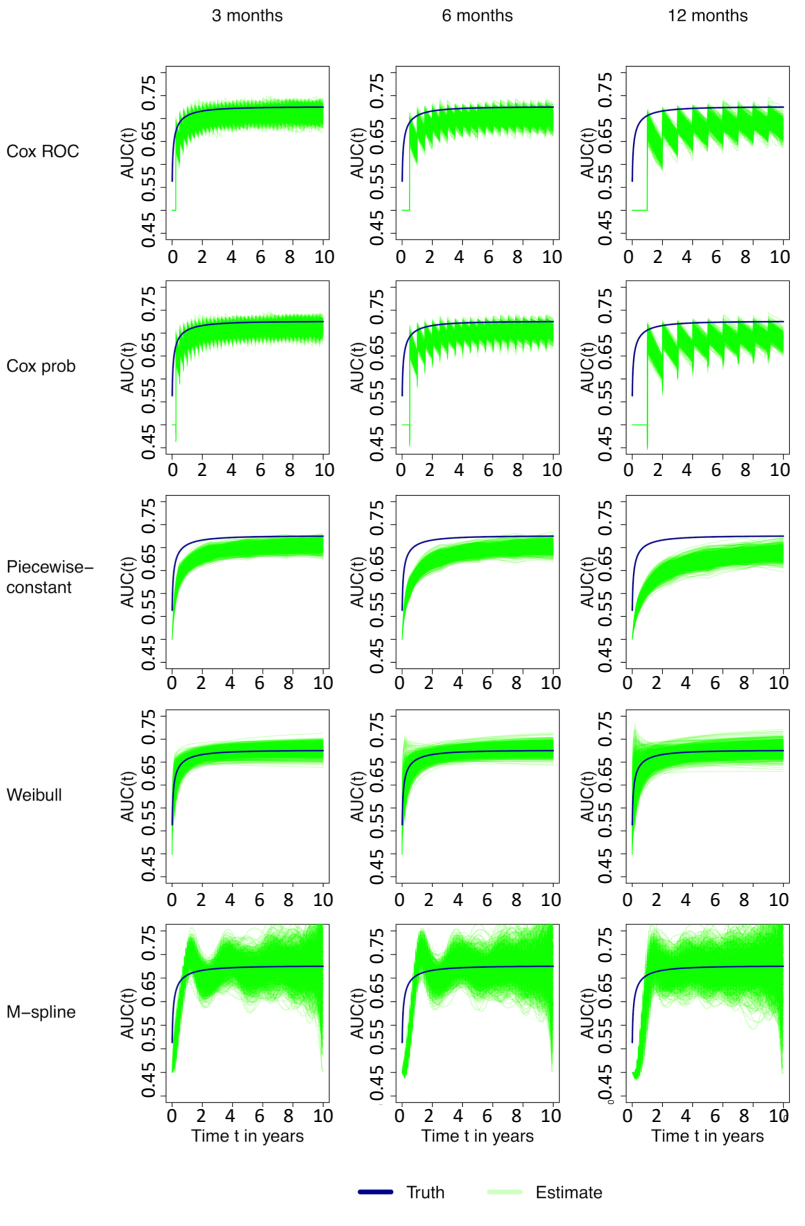

Figure 3.C.3: Incident/dynamic AUC for scenario J (3 months), K (6 months) and L (12 months). ***Abbreviations:*** *Cox ROC, estimate based on* `risksetAUC` *function; Cox prob, estimate based on transition probabilities of Cox model.*

§3.D Cumulative/dynamic AUC for scenarios D–L

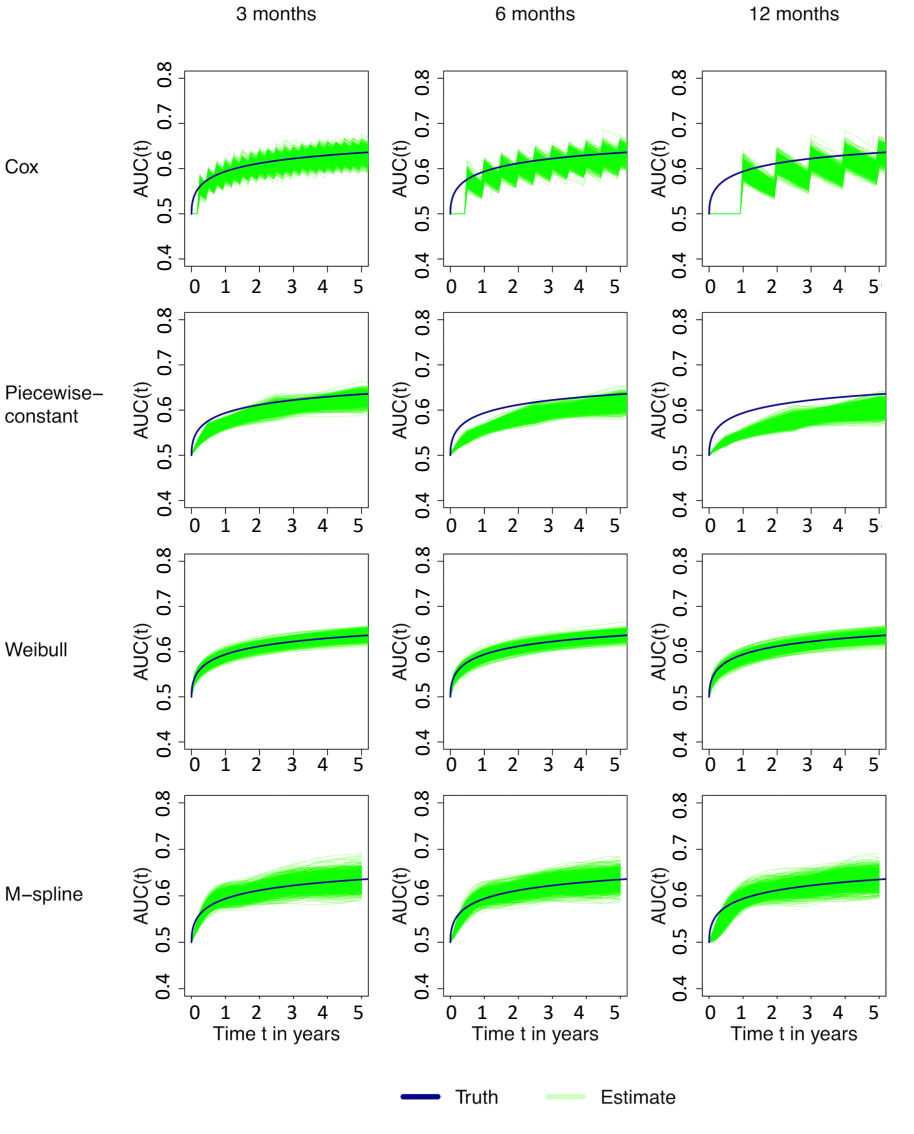

Figure 3.D.1: Cumulative/dynamic AUC for scenario D (3 months), E (6 months) and F (12 months).

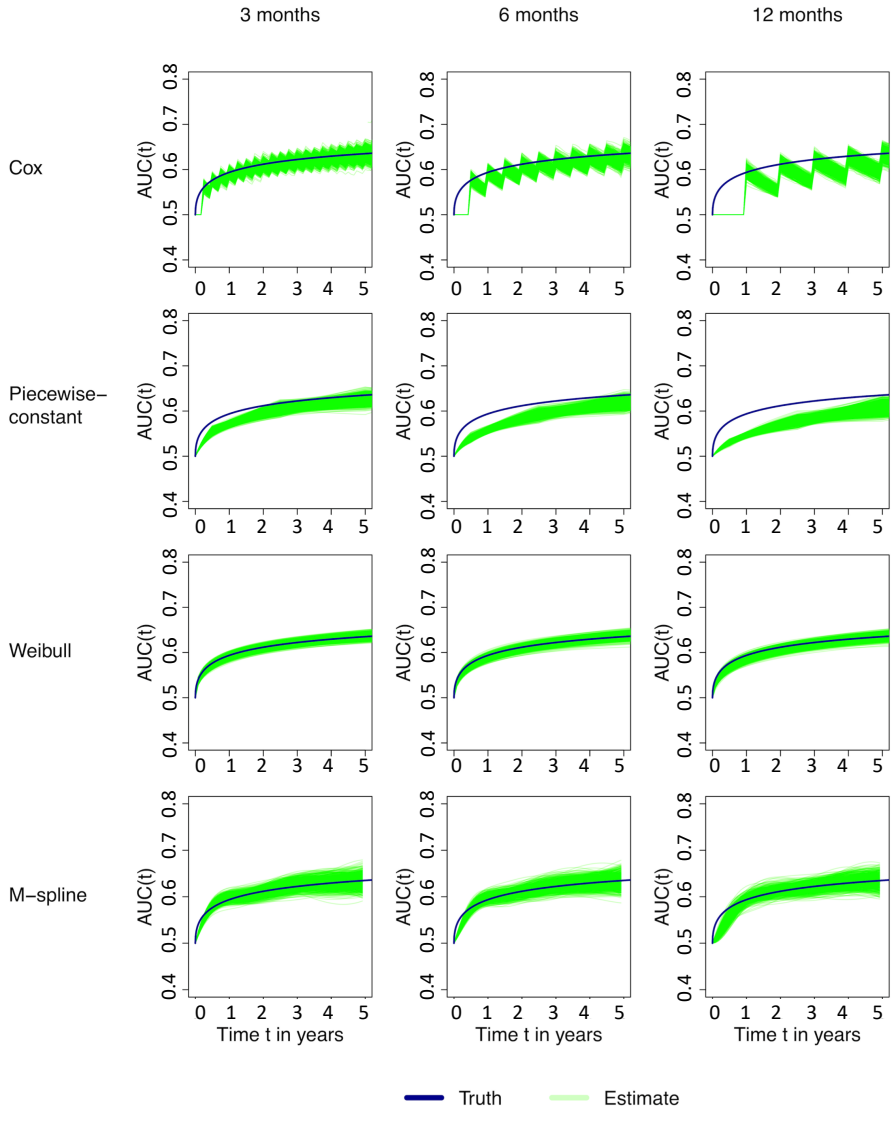

Figure 3.D.2: Cumulative/dynamic AUC for scenario G (3 months), H (6 months) and I (12 months).

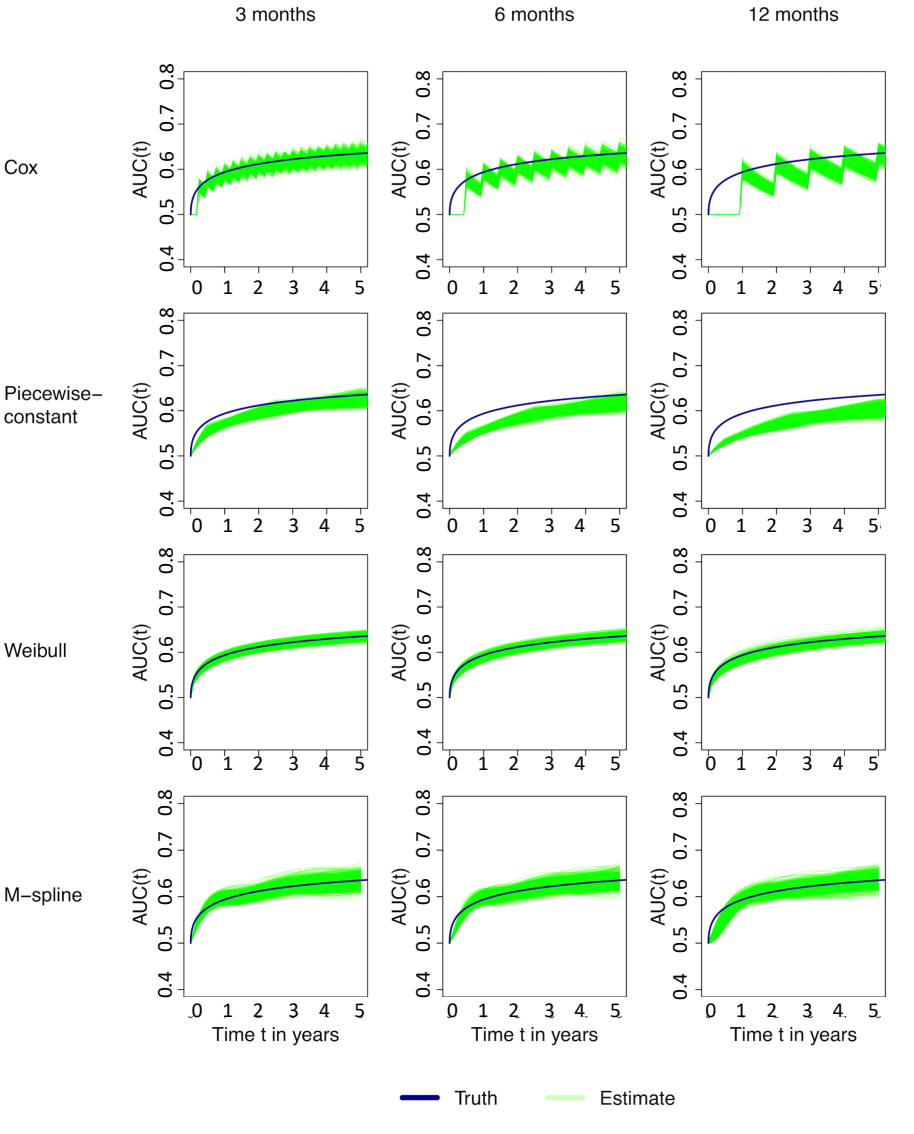

Figure 3.D.3: Cumulative/dynamic AUC for scenario J (3 months), K (6 months) and L (12 months).