Andreas Kuehnapfel¹ / Fabian Schwarzenberger¹ / Markus Scholz¹

On the Conditional Power in Survival Time Analysis Considering Cure Fractions

¹ Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig, Haertelstrasse 16-18, 04107 Leipzig, Germany, E-mail: andreas.kuehnapfel@imise.uni-leipzig.de

Abstract:

Conditional power of survival endpoints at interim analyses can support decisions on continuing a trial or stopping it for futility. When a cure fraction becomes apparent, conditional power cannot be calculated accurately using simple survival models, e.g. the exponential model. Non-mixture models consider such cure fractions. In this paper, we derive conditional power functions for non-mixture models, namely the non-mixture exponential, the non-mixture Weibull, and the non-mixture Gamma models. Formulae were implemented in the R package CP. For an example data set of a clinical trial, we calculated conditional power under the non-mixture models and compared results with those under the simple exponential model.

Keywords: cure fraction, exponential survival, Gamma type survival, non-mixture model, proportional hazards, randomised clinical trial, Weibull type survival

DOI: 10.1515/ijb-2015-0073

1 Introduction

Clinical trials often consider survival endpoints which occur relatively late compared to treatment. Interim analyses are common practice to decide early whether to stop the trial for ethical reasons or futility. Conditional power is an important concept in this decision-making process. Conditional power is defined as the probability of achieving a significant result at the end of the study in case of an existing effect, given data from interim analysis. A commonly applied simple approach to estimate (conditional) power is to use an exponential survival model. Here, the hazard is constant and survival tends to zero as time tends to infinity. However, there are several clinically relevant questions where this assumption is not appropriate. For example, cure or relapse of disease might be of interest. Here, not all individuals experience an event resulting in a plateau of the survival curves. Other non-medical examples are discharge of prisoners or malfunction of a particular component of a car. For substantial cure fractions it is clear that the simple exponential model is not appropriate to explain the data. Lambert [1] gives a good overview of cure fraction models which are better suited for that situation, e.g. the so-called non-mixture models. Conditional power formulae for the simple exponential model are already derived by Andersen [2]. For cure fraction models, no such formulae are available so far. In the present paper, we aim at closing this gap: In Section 2, a summary for the exponential model is provided including corresponding unconditional and conditional power estimates. Section 3 introduces the non-mixture models. The corresponding unconditional and conditional power functions are derived for the non-mixture exponential model under the proportional hazard assumption. We provide a simulation study to interpret the unconditional power estimates and to compare them with corresponding conditional estimates. We also provide a simulation study to empirically validate conditional power estimates for selected scenarios. Finally, we consider alternative parametric assumptions for baseline hazard, namely non-mixture Weibull and non-mixture Gamma models. In Section 4, the conditional power formulae will be applied to a data set of a randomised clinical trial. The R package CP containing all formulae derived in this paper is presented in brief. Results are discussed in Section 5. Theorems and some auxiliary results can be found in the Appendix.

2 Subjects of interest

In this section, we recall and describe the survival models studied in the present paper. We restrict our considerations to comparisons of two treatment groups (1 and 2) throughout. First, we consider the exponential model

Andreas Kuehnapfel is the corresponding author.

© 2017 Walter de Gruyter GmbH, Berlin/Boston.

This content is free.

and corresponding unconditional and conditional power estimates. Then, we introduce survival fractions and specify the class of models considered here.

2.1 The exponential survival model

In the simplest way, survival S_i of group i at time t can be modelled using the exponential distribution with rate parameter $\lambda_i > 0$, i.e.

$$S_i(t) = \exp(-\lambda_i t),$$

with the constant hazard

$$h_i(t) = \lambda_i.$$

The hazard ratio is defined as

$$\theta = \frac{\lambda_2}{\lambda_1}.$$

For superiority trials, the test hypotheses are

$$H_0: \theta = 1 \quad \text{vs.} \quad H_1: \theta \neq 1.$$

The parameter λ_i can be estimated by the maximum likelihood method: Let D_i be the number of deaths in group i and O_i be the total observation time of group i one obtains

$$\hat{\lambda}_i = \frac{D_i}{O_i},$$

and with it, the hazard ratio can be estimated as

$$\hat{\theta} = \frac{\hat{\lambda}_2}{\hat{\lambda}_1}.$$

Examples for the exponential survival model can be found in Figure 1.

2.2 Unconditional power based on the exponential model

According to Andersen [2], one way of defining an appropriate test statistic of the difference in survival is

$$W = \log \hat{\theta} = \log \hat{\lambda}_2 - \log \hat{\lambda}_1 = \log \frac{D_2}{O_2} - \log \frac{D_1}{O_1}. \quad (1)$$

With Theorem 1 in Section A of the Appendix, one can define the asymptotic power function by

$$\gamma(\theta) = \Phi\left(\frac{q_{\frac{\alpha}{2}} \sigma_{H_0} + \mu_{H_0} - \mu_{H_1}(\theta)}{\sigma_{H_1}}\right) + \left\{1 - \Phi\left(\frac{q_{1-\frac{\alpha}{2}} \sigma_{H_0} + \mu_{H_0} - \mu_{H_1}(\theta)}{\sigma_{H_1}}\right)\right\} \quad (2)$$

where Φ is the cumulative distribution function of the standard normal distribution and q_α is its α -quantile. The mean and variance under the null are specified by $\mu_{H_0} = 0$ and $\sigma_{H_0}^2 = 2/n$ where n is the required sample size. For mean and variance under the alternative it holds $\mu_{H_1}(\theta) = \log \theta$ and $\sigma_{H_1}^2 = 2/n$.

2.3 Conditional power based on the exponential model

Suppose that data $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$ at time of interim analysis is available. Let D_i^* be the number of future deaths and O_i^* be the future total observation time of group i measured from interim analysis until the end of study. Referring to the unconditional test statistic eq. (1) derived in Section 2.2 one obtains

$$W^* = \log \frac{D_2 + D_2^*}{O_2 + O_2^*} - \log \frac{D_1 + D_1^*}{O_1 + O_1^*}$$

as the conditional test statistic.

Using Theorem 2 in Section A of the Appendix, the asymptotic conditional power function given data $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$, reads as follows

$$\gamma_c(\theta) = \Phi\left(\frac{q_{\frac{\alpha}{2}}\sigma_{H_0} + \mu_{H_0} - \mu_{H_1}(\theta)}{\sigma_{H_1}}\right) + \left\{1 - \Phi\left(\frac{q_{1-\frac{\alpha}{2}}\sigma_{H_0} + \mu_{H_0} - \mu_{H_1}(\theta)}{\sigma_{H_1}}\right)\right\}, \quad (3)$$

where means and variances under the null and under the alternative are

$$\begin{aligned} \mu_{H_0} &= \log\left(\frac{d_2 + \lambda_1 EO_2^*}{o_2 + EO_2^*}\right) - \log\left(\frac{d_1 + \lambda_1 EO_1^*}{o_1 + EO_1^*}\right), \\ \sigma_{H_0}^2 &= \frac{2}{n^*}, \\ \mu_{H_1}(\theta) &= \log\left(\frac{d_2 + \theta\lambda_1 EO_2^*}{o_2 + EO_2^*}\right) - \log\left(\frac{d_1 + \lambda_1 EO_1^*}{o_1 + EO_1^*}\right) \\ \sigma_{H_1}^2 &= \frac{2}{n^*}. \end{aligned}$$

Here, n^* denotes the remaining sample size, i.e. the total sample size n minus the number of patients included so far. E is the operator for the expected value.

It should be mentioned that the formulae derived here are not exactly the same as those proposed by Andersen [2]. The main difference is our centering of the asymptotic conditional power function (3) on μ_{H_0} which is generally not equal to zero in contrast to the unconditional case.

2.4 Survival models with cure fractions

If a relevant cure fraction is observed, the application of the exponential survival model considered above is no longer appropriate. In the literature, two parametric models dealing with cure fractions are most frequently in use: mixture and non-mixture models, cf. Martinez et al. [3]. We favour parametric models due to the availability of maximum likelihood methods. They are easier to interpret and provide higher power in statistical testing than comparable non-parametric models. The mixture model can be formulated as the sum of the proportion of the survivors and the weighted survival of the non-survivors [1]:

$$S_i(t) = c_i + (1 - c_i)S_i^0(t)$$

with a cure fraction $c_i \in (0, 1)$ and some ordinary survival function S_i^0 . Thus, one has $S_i(0) = 1$ and $S_i(t) \rightarrow c_i$ as $t \rightarrow \infty$.

With the same notation and limit behaviour, the improper survival function of the non-mixture model reads as follows [1]:

$$S_i(t) = c_i^{1-S_i^0(t)}.$$

In the present paper, we restrict considerations on non-mixture models. They require the specification of proper survival functions S_i^0 . We assume in the following that S_i^0 are from the exponential, the Weibull, or the Gamma class but consider the exponential class first. It will turn out that for an explicit derivation of power formulae, it is essential to assume proportional hazards for the two treatment groups to be compared. In terms of the non-mixture model, this implies that $S_1^0 = S_2^0$, i.e. rate and/or shape parameters of the above mentioned parametric distributions are assumed to be equal for the two treatment groups, e.g. $\lambda_1 = \lambda_2$ for the exponential class. Thus, the hazard ratio can be expressed as follows:

$$\theta = \frac{\log c_2}{\log c_1}, \quad (4)$$

i.e. the hazard ratio is the ratio of the log-cure fractions. Further considerations and discussions of the non-mixture model can be found in Lambert [1] and Tsodikov et al. [4, 5]. Examples of the parametric survival models considered here can be found in Figure 1.

Figure 1: Examples of the four parametric survival models considered here. For the non-mixture models, we assumed a survival fraction of 0.3 throughout this figure.

3 The non-mixture model with exponential survival

In this section, we derive the unconditional and conditional power formulae for the non-mixture exponential model. We perform a simulation study to illustrate the results of the unconditional power analysis and to compare them with those of the exponential model without assuming cure fractions. To validate our findings, we study the agreement of our conditional power estimates with empirical power estimates for selected scenarios. Finally, we consider parametric alternatives to the non-mixture exponential model.

3.1 Basic requirements

If S_i^0 are from the exponential class, one has to specify the rate parameters $\lambda_i > 0$, i.e. $S_i^0(t) = \exp(-\lambda_i t)$. The corresponding hazard functions h_i for S_i are then

$$h_i(t) = -\lambda_i \exp(-\lambda_i t) \log c_i \quad (5)$$

(see Lemma 1 in Section B of the Appendix) which yields the hazard ratio

$$\theta(t) = \frac{\lambda_2}{\lambda_1} \exp\{-(\lambda_2 - \lambda_1)t\} \frac{\log c_2}{\log c_1}.$$

Thus time independence of the hazard ratio can only be achieved if

$$\lambda_1 = \lambda_2$$

resulting in eq. (4). For clarification, this does not mean that we assume constant hazards (cf. eq. (5)) as it would be the case for the simple exponential model. Again, we consider the hypotheses

$$H_0: \theta = 1 \quad \text{vs.} \quad H_1: \theta \neq 1.$$

An estimator of the hazard ratio under $\lambda_1 = \lambda_2$ can be defined by

$$\hat{\theta} = \frac{\log \hat{c}_2}{\log \hat{c}_1}$$

(see Lemma 2 in Section B of the Appendix for more details).

3.2 Power estimate for the unconditional case

Similarly to eq. (1) in Section 2.2, one can start with

$$W = \log \hat{\theta} = \log(-\log \hat{c}_2) - \log(-\log \hat{c}_1) = \log \frac{D_2}{O'_2} - \log \frac{D_1}{O'_1} \quad (6)$$

as the test statistic of interest, with $D_i = \sum_{j_i=1}^{n_i} \Delta_{i,j_i}$ and $O'_i = \sum_{j_i=1}^{n_i} \{1 - \exp(-\hat{\lambda}_i Y_{i,j_i})\}$. Here, Δ_{i,j_i} for patient j_i of group i indicates whether the patient is dead ($\Delta_{i,j_i} = 1$) or censored ($\Delta_{i,j_i} = 0$) whereas Y_{i,j_i} denotes the time of event or censoring. Using Theorem 3 in Section A of the Appendix, one can define the asymptotic power function via eq. (2) of Section 2.2. Here, the mean and variance under the null are given by $\mu_{H_0} = 0$ and $\sigma_{H_0}^2 = 2[n\{1 - c_1(1 + \log^2 c_1)\}]^{-1}$ whereas under the alternative it holds that $\mu_{H_1}(\theta) = \log \theta$ and $\sigma_{H_1}^2(\theta) = [n\{1 - c_1(1 + \log^2 c_1)\}]^{-1} + [n\{1 - c_1^\theta(1 + (\theta \log c_1)^2)\}]^{-1}$.

3.3 Power estimate for the conditional case

Again, we consider the situation that data $(D_1, O'_1, D_2, O'_2) = (d_1, o'_1, d_2, o'_2)$ is available from interim analysis. Let D_i^* denote the number of deaths and let O_i^{*} be a function of the future observation time of group i counted from time of interim analysis until final analysis of the study in analogy to the definition of O'_i at the beginning of Section 3.2. The conditional test statistic can be defined in analogy to eq. (6) in Section 3.2 by

$$W^* = \log \frac{D_2 + D_2^*}{O'_2 + O_2^{*}} - \log \frac{D_1 + D_1^*}{O'_1 + O_1^{*}}.$$

With the help of Theorem 4 in Section A of the Appendix, the asymptotic conditional power function given $(D_1, O'_1, D_2, O'_2) = (d_1, o'_1, d_2, o'_2)$, can be derived. Using eq. (3) of Section 2.3, corresponding means and variances under the null and the alternative are given by

$$\begin{aligned} \mu_{H_0} &= \log \left(\frac{d_2 - \log(c_1)EO_2^{*}}{o'_2 + EO_2^{*}} \right) - \log \left(\frac{d_1 - \log(c_1)EO_1^{*}}{o'_1 + EO_1^{*}} \right), \\ \sigma_{H_0}^2 &= \left[n^* \left\{ 1 - \exp \left(-\frac{d_1 - \log(c_1)EO_1^{*}}{o'_1 + EO_1^{*}} \right) \left(1 + \left(\frac{d_1 - \log(c_1)EO_1^{*}}{o'_1 + EO_1^{*}} \right)^2 \right) \right\} \right]^{-1} \\ &\quad + \left[n^* \left\{ 1 - \exp \left(-\frac{d_2 - \log(c_1)EO_2^{*}}{o'_2 + EO_2^{*}} \right) \left(1 + \left(\frac{d_2 - \log(c_1)EO_2^{*}}{o'_2 + EO_2^{*}} \right)^2 \right) \right\} \right]^{-1}, \\ \mu_{H_1}(\theta) &= \log \left(\frac{d_2 - \theta \log(c_1)EO_2^{*}}{o'_2 + EO_2^{*}} \right) - \log \left(\frac{d_1 - \log(c_1)EO_1^{*}}{o'_1 + EO_1^{*}} \right) \\ \sigma_{H_1}^2(\theta) &= \left[n^* \left\{ 1 - \exp \left(-\frac{d_1 - \log(c_1)EO_1^{*}}{o'_1 + EO_1^{*}} \right) \left(1 + \left(\frac{d_1 - \log(c_1)EO_1^{*}}{o'_1 + EO_1^{*}} \right)^2 \right) \right\} \right]^{-1} \\ &\quad + \left[n^* \left\{ 1 - \exp \left(-\frac{d_2 - \theta \log(c_1)EO_2^{*}}{o'_2 + EO_2^{*}} \right) \left(1 + \left(\frac{d_2 - \theta \log(c_1)EO_2^{*}}{o'_2 + EO_2^{*}} \right)^2 \right) \right\} \right]^{-1}. \end{aligned}$$

3.4 Interpretation and comparison of unconditional power estimates

Here, we compare unconditional power estimates under the non-mixture exponential model with those of the simple exponential model. To illustrate the difference between the models, we choose $\alpha = 5\%$ and select three scenarios for (unconditional) power calculation. Sample size is kept fixed and was calculated to achieve a power of 80% for an initially chosen hazard ratio θ . The formula of Andersen [2] was used for this purpose. Power of the exponential and the non-exponential model were then calculated in dependence on modified hazard ratios. We considered the following scenarios:

Scenario 1: We fix the baseline survival rate $c_1 = 30\%$ and choose an initial hazard ratio $\theta = 0.872$, resulting in a survival rate $c_2 = 35\%$ for the experimental arm. According to Andersen, $n = 2481$ patients are required for 80% power. Keeping the sample size constant and modifying the hazard ratio results in Figure 2 (left panel).

Scenario 2: We set the survival rate of the standard arm to $c_1 = 40\%$ and choose an initial hazard ratio of $\theta = 0.756$, resulting in a survival rate of $c_2 = 50\%$ for the experimental arm. Thus, simulations are performed for $n = 738$ patients. The corresponding relation of hazard ratio and power is displayed in Figure 2 (mid panel).

Scenario 3: We assume a baseline survival rate $c_1 = 50\%$ and set the initial hazard ratio to $\theta = 0.621$, resulting in a survival rate $c_2 = 65\%$ for the experimental arm. Simulations are performed for $n = 337$ patients. The resulting power plot is displayed in Figure 2 (right panel).

Assuming identical sample size and the same hazard ratio, power estimates under the non-mixture model are typically lower than those under the simple exponential model. Keeping the hazard ratio constant, the difference increases with baseline survival (cf. Figure 2). Survival rates can be chosen as one-, five-, or ten-year survival rates, for example, depending on the question of interest. These considerations also hold for the other non-mixture models considered here (non-mixture Weibull and non-mixture Gamma model, see below). In the situation of an interim analysis, available data might raise doubts regarding the survival model originally assumed. This might result in a revision of model parameters resulting in new power estimates. As expected, for both, exponential and non-mixture exponential models, power drops if the effect size is smaller than expected (see Figure 2).

Figure 2: Power for the exponential (black) and the non-mixture exponential model (red) for three different scenarios. Left panel: $c_1 = 30\%$, $n = 2481$. Mid panel: $c_1 = 40\%$, $n = 738$. Right panel: $c_1 = 50\%$, $n = 337$.

Conversely, our formulae can be used to re-calculate the sample sizes for the scenarios considered in Figure 2. Results are shown in Figure 3.

Figure 3: Sample size estimation of the scenarios considered in Figure 2. Sample size is calculated to achieve a power of 80%. Left panel: $c_1 = 30\%$. Mid panel: $c_1 = 40\%$. Right panel: $c_1 = 50\%$.

One can also interpret these results in terms of “information fraction”. Here, information fraction is defined as the ratio of the observed number of events at a given time point to the expected number of events at the end of the study. Then, the question is whether there are differences between the information fraction given

a simple exponential model or given a non-mixture model. As mentioned above, event rates differ between both model types. The simple exponential model has a constant hazard whereas the hazard of the non-mixture model changes over time, being high at the early phase of the trial and decreasing to zero in the later phase of the trial. Thus, information fraction of an underlying non-mixture model will be higher at earlier time points and lower at later time points of a study compared to the information fraction under a simple exponential model at the same time points.

3.5 Empirical accuracy of the conditional power estimates

In this section, we illustrate accuracy of our conditional power formula for the non-mixture exponential model by performing a simulation study. The simulation study is implemented in the following way. First, we specify the parameters c_i and λ_i of our non-mixture exponential model, and with it, the hazard ratio θ . Then, we simulate data for a fictive five-year clinical trial based on the above specified survival model under our proportional hazard assumption, i.e. $\lambda_1 = \lambda_2$. Subject recruitment is distributed equally over study duration. We then simulate an interim analysis after 2.5 years by censoring events occurring after the time point of interim analysis. Interim data is used to calculate conditional power assuming a non-mixture exponential model. Further recruitment of subjects is considered within conditional power formula (equally until the end of study). For comparison, we apply the Wald test to the entire data set to decide whether the difference in survival at the end of the study is significant (value 1) or not (value 0). We repeat this simulation procedure 1,000 times and compare the average conditional power with the average of the Wald test decisions which can be interpreted as an empirical power. The following scenarios are considered:

Scenario 1: We choose a baseline survival rate $c_1 = 70\%$ and a hazard ratio $\theta = 0.626$, resulting in a survival rate $c_2 = 80\%$ for the experimental arm. The rate parameters are fixed at $\lambda_1 = \lambda_2 = 0.1$. We perform simulations for $n = 500$ patients. Conditional power by the non-mixture exponential model yields 32%, the percentage of significant Wald tests was the same (32%).

Scenario 2: We set the survival rate of the standard arm to $c_1 = 65\%$ and choose a hazard ratio $\theta = 0.377$, resulting in a survival rate of $c_2 = 85\%$ for the experimental arm. Finally, we choose $\lambda_1 = \lambda_2 = 0.2$. Simulations are performed for $n = 250$ patients. Conditional power based on the non-mixture exponential model is 67% while 70% of the Wald tests were significant.

Scenario 3: We assume a baseline survival rate $c_1 = 60\%$ and a hazard ratio $\theta = 0.318$, resulting in a survival rate $c_2 = 85\%$ for the experimental arm. The rate parameters are set to $\lambda_1 = \lambda_2 = 0.5$. Simulations were performed for $n = 150$ patients. Conditional power of the non-mixture exponential model yields 62% whereas 59% of the Wald tests are significant.

As one can see, conditional power estimates and Wald test results are in good agreement for the selected scenarios. To facilitate further simulations, we provide the corresponding R script as supplement material.

3.6 Alternative parametric assumptions for baseline hazard

The non-mixture exponential model might not be flexible enough to model the data adequately. Two possible alternatives are considered here: the non-mixture Weibull model and the non-mixture Gamma model. We introduce their notation now. The derivation of the (conditional) power functions is similar to that of the non-mixture exponential model. Results are presented in the Appendix.

3.6.1 Weibull-type survival

The survival function of group i according to the Weibull distribution is of the form

$$S_i^0(t) = \exp(-\lambda_i t^{k_i})$$

with scale parameter $\lambda_i > 0$ and shape parameter $k_i > 0$. For maximum likelihood estimates see Remark 1 in Section B of the Appendix.

Examples for the non-mixture Weibull model can be found in Figure 1.

For a derivation of the conditional power formula, we refer to Section C.2 of the Appendix.

3.6.2 Gamma-type survival

Another option to increase flexibility is to model survival functions with the Gamma distribution. The main difference to the non-mixture Weibull model is a different shape of the hazard function (see Remark 3 in Section B of the Appendix).

The incomplete Gamma function of the upper bound is defined by

$$\gamma(x, y) = \int_0^y s^{x-1} \exp(-s) ds$$

and the Gamma function is given by

$$\Gamma(x) = \int_0^{\infty} s^{x-1} \exp(-s) ds.$$

The regularised incomplete Gamma function of the upper bound is defined by

$$\Gamma^0(x, y) = \frac{\gamma(x, y)}{\Gamma(x)}.$$

Then, the survival function of group i is

$$S_i^0(t) = 1 - \Gamma^0(a_i, b_i t) = \frac{\int_0^{\infty} s^{a_i-1} \exp(-s) ds}{\int_0^{\infty} s^{a_i-1} \exp(-s) ds}$$

with shape parameter $a_i > 0$ and rate parameter $b_i > 0$. Maximum likelihood estimates are given in Remark 2 in Section B of the Appendix.

Examples for the non-mixture Gamma model can be found in Figure 1.

We refer to Section C.3 of the Appendix for a derivation of the conditional power formula.

4 Application

The unconditional and conditional power formulae for the exponential and the non-mixture models were implemented in the R package CP which we describe in the following. A function for the comparison of these models is also provided. To demonstrate the models and to illustrate their differences, we apply the package to a data set of a clinical trial for which a futility analysis has been performed during interim analysis.

4.1 The R package CP

We implemented the formulae derived in this paper in the framework of the R package CP which can be downloaded via CRAN. The function `ConPwrExp` contains the formulae for the exponential model without any cure fraction while the functions for the non-mixture models are `ConPwrNonMixExp` (exponential survival), `ConPwrNonMixWei` (Weibull-type survival) and `ConPwrNonMixGamma` (Gamma-type survival), respectively. A function for the comparison of these four models on the basis of Akaike's information criterion (AIC) [6] is also included (`CompSurvMod`). Finally, conditional power according to Andersen (`ConPwrExpAndersen`) [2] can be calculated which is also based on an exponential model. Output comprises a summary of survival data of the interim analysis, AICs, parameter estimates, and estimates for the conditional power for the chosen model. Plots of the Kaplan-Meier curves and estimated parametric survival curves are provided. Further details and explanation of package parameters are presented in Table 1 and are available via the help pages of CP.

Table 1: Description of parameters of the R package CP.

Parameter	Description	Default
-----------	-------------	---------

data	Data frame which consists of at least three columns with the group in the first, survival status in the second and event time in the third column	
cont.time	Trial duration after interim analysis	
new.pat	two-dimensional vector which consists of numbers of new patients being recruited at each time unit and corresponding study arm	(0, 0)
theta.0	Clinically relevant difference assumed for initial power planning	1
alpha	Significance level for conditional power calculations	0.05
disp.data	Logical value indicating whether all results of calculations should be displayed	FALSE
plot.km	Logical value indicating whether Kaplan-Meier curves and estimated survival curves assuming the exponential model should be plotted	FALSE

4.2 Clinical example

We applied our methods to an interim analysis of the highCHOEP trial of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL) [7]. The question to be answered by this trial is whether there is a difference in survival between patients treated with either standard dosage chemotherapy (CHOEP) or dose-intensified chemotherapy (highCHOEP). Patients were between 18 and 60 years old and randomly received one of the above mentioned therapies as first-line treatment. At time of interim analysis, a total of 233 (CHOEP: 118, highCHOEP: 115) patients were analysed, revealing 69 (CHOEP: 33, highCHOEP: 36) death events and an aggregated observation time of 4306 (CHOEP: 2191, highCHOEP: 2115) months (see also Table 2).

Table 2: Data used for interim analysis: CHOEP is the standard-dose chemotherapy, highCHOEP is the experimental arm with dose-escalated therapy.

	CHOEP	highCHOEP	Total
Patients	118	115	233
Deaths	33	36	69
Censored	85	79	164
Person months	2,191	2,115	4,306

The initial study was planned under an exponential survival model without cure fractions. A hazard ratio of 0.653 was assumed to be clinically relevant. We calculate conditional power for a total sample size of 389 patients assuming an accrual rate of 5.5 patients per month and group resulting in a final analysis after 14.2 months.

The function `CompSurvMod` of our R-package CP can now be used to calculate the conditional power for each of the models. Note that the hazard ratio translates into a ratio of log-cure fractions within our non-mixture models assuming proportional hazards. AIC can be used to decide which of the four models fits best. Lower AIC represents the better compromise between model complexity and fit. Calculated likelihoods, AICs, estimated parameters for survival functions and expected future observation times are shown in Table 3.

Table 3: Comparison of four models: We present AICs, sums of AICs of both arms, likelihoods, parameter estimates and future person months for each model. As can be seen from the AIC values, all non-mixture models clearly fit the data better than the exponential model. The non-mixture Weibull model performed the best.

Model		CHOEP	highCHOEP
Exponential	Likelihood	− 171.4636	− 182.6507
$\sum$ AIC = 712.2286	AIC	344.9271	367.3015
$\hat{\theta} = 1.1300$	$\hat{\lambda}$	0.0151	0.0170
	O^*	1,603	1,508
Non-mixture-exponential	Likelihood	− 168.3726	− 179.0935
$\sum$ AIC = 702.9322	AIC	340.7452	362.1870
$\hat{\theta} = 1.1352$	$\hat{\lambda}$	0.0577	0.0577
	$\hat{c}$	0.6267	0.5883
	O^*	1,553	1,455
Non-mixture-Weibull	Likelihood	− 167.4513	− 179.5084
$\sum$ AIC = 701.9195	AIC	338.9027	363.0168
$\hat{\theta} = 1.1348$	$\hat{\lambda}$	0.0480	0.0480
	$\hat{k}$	1.1482	1.1482
	$\hat{c}$	0.6569	0.6207

Non-mixture-Gamma $\sum \text{AIC} = 705.9129$	O^*	1,555	1,455
	Likelihood	– 166.4837	– 182.4728
	AIC	336.9673	368.9456
	$\hat{a}$	1.6495	1.6495
$\hat{\theta} = 1.1334$	$\hat{b}$	0.1404	0.1404
	$\hat{c}$	0.6721	0.6374
	O^*	1,554	1,455

As one can see in Table 3, at time of interim analysis the estimated hazard ratios $\hat{\theta}$ are close to 1.1 which is far away from the expected $\theta = 0.653$. With respect to the AIC, the non-mixture model with Weibull-type survival shows the best fit among the four models considered. Table 3 also shows, that compared to the exponential model, all three non-mixture models make a better compromise between data fitting and the number of additional parameters to be estimated. Moreover, the proportional hazard assumption for the non-mixture models, i.e. the equality of the rate, scale, and shape parameters of both groups is justified (Non-mixture-exponential: $p = 0.9995$, Non-mixture-Weibull: $p = 0.3921$, Non-mixture-Gamma: $p = 0.1337$, Likelihood ratio test). Figure 4 illustrates the Kaplan-Meier survival estimates of the interim data in comparison to the estimated parametric survival functions. Again, it can be recognised that the simple exponential model does not fit the data very well, in contrast to the non-mixture models. We present the conditional power curves in Figure 5, i.e. the dependence of power on the true effect size for the four models. The conditional power estimates for the planned effect ($\theta = 0.653$) are: 39.2% for the exponential, 18.6% for the non-mixture exponential, 14.2% for the non-mixture Weibull, and 17.1% for the non-mixture Gamma model. Thus, a considerable difference in power estimates is obtained when switching to non-mixture models. The exponential model clearly over-estimates the conditional power due to over-estimation of the number of additional events. In contrast, differences in power estimates between the non-mixture models are relatively small. For interpreting these conditional power estimates it is necessary to provide further details regarding original study planning. The study was planned under the exponential model with 670 patients to achieve a power of 82%. Total sample size accounts for 10% of patients lost to follow-up or missing data. Due to an amendment prescribing additional treatment with the antibody Rituximab for every newly recruited patient, study analysis was performed for the 389 patients recruited thus far. Using the formula of Andersen, the reduced number of individuals results in a power of 59%. Performing the same calculations using the non-mixture exponential model results in a power of 95% for the 670 originally planned patients and 79% for the 389 available patients. Remembering our comparison of unconditional power estimates in Section 3.4, these results can be explained by two concurrent mechanisms: First, considering cure fractions results in fewer events and thus lower power in the longer perspective. However, events are generated more quickly under the non-mixture exponential model which could result in improved power at a shorter perspective. In other words, the survival curves assuming a simple exponential versus a non-mixture exponential model are typically intersecting if fitted to the same data (see also Figure 4). Thus, it is difficult to predict, which model results in the higher predicted power at a specified time point of final data analysis.

Figure 4: Comparison of parametric models and survival data: While the simple exponential model does not fit the data very well, the non-mixture models result in reasonable fits.

Figure 5: Conditional power functions: We present conditional power estimates in dependence on the true effect size for each of the models.

5 Discussion

5.1 Motivation of the work

Large clinical trials with survival endpoints are often time-consuming and costly. Therefore, interim analyses are common practice. If estimated effect size at time of interim analysis is far below the clinical expectation, conditional power might be a useful aid to evaluate futility of further recruitment which can result in premature termination of the trial. In case of a significant proportion of patients never experiencing an event, simple exponential survival models as proposed by Andersen [2] do not adequately describe the data. In the present

paper, we derive conditional power formulae for three non-mixture models assuming cure fractions, namely those of exponential, Weibull, or Gamma type. We restricted our considerations to comparisons of two study arms and assume proportional hazards throughout the paper. Conditional power functions were implemented in an R package and are applied to a clinical data set.

5.2 Modelling Survival Data with Cure Fractions

There are three common types of modelling survival with cure fractions: The parametric, the semi-parametric, and the non-parametric approach, each of them having its own advantages and disadvantages [8]. Since the parametric approach offers a convenient way of testing proportional hazard assumptions and applying maximum likelihood methods for estimation purposes, this kind of modelling is chosen here. Asymptotic normality and unbiasedness of estimates in combination with the delta method allowed us to derive asymptotic distributions of the test statistic considered, which in our case, is the logarithm of the estimator of the hazard ratio.

Fitting piecewise exponential models could be an interesting alternative since it covers both cases, with and without cure fraction. It might also be appropriate for situations for which different event mechanisms could be expected, e.g. during and after an intense therapy. However, we believe that there are also disadvantages to this approach: First, the proportional hazard assumption is less obvious for this setting, especially if there are different event mechanisms to be considered. Without this assumption, the test statistic used in this paper is time dependent complicating all conditional power considerations. Second, the intervals of constant hazards need to be chosen and need to be identical for the treatment arms. This increases the number of free parameters to be determined or requires additional assumptions. Finally, conditional power relies on extrapolation of hazards. Considering a model of piece-wise constant hazards requires assumptions regarding the future of the piecewise behavior. Estimation of hazard ratios differ from exponential to non-mixture models. However, the relevant test statistic is the same. A problem could arise for non-mixture models if the regularity conditions required for asymptotic properties of maximum likelihood estimators are violated. This could occur due to the usage of quasi-probability density functions, i.e. survival functions bounded away from zero. Here, these regularity conditions are fulfilled (proof not shown).

Another approach of modelling survival is provided by the so-called mixture models. These models often give results very similar to those of the non-mixture models [9]. Further details can be found in Boag [10], Berkson and Gage [11], Sposto and Sather [12], and Sposto [13]. A completely different approach to estimating conditional power is via Bayesian methodology. This approach circumvents the issue of asymptotics, is often easier to calculate, and is directly interpretable. However, it faces the problem of specifying appropriate priors [14, 15].

5.3 Proportional Hazard Assumption

We assumed proportional hazards of the two study arms to be compared. In this sense, the non-mixture models as well as the exponential model are related to the Cox model with specified baseline hazard [16]; see Section 2.1 for the exponential model, eq. (5) for the non-mixture model with exponential survival, and Remark 3 for the non-mixture models with Weibull-type survival and Gamma-type survival. For all non-mixture models, proportional hazards imply that the hazard ratio equals the ratio of log-cure fractions. It allows applying the delta method to derive explicit formulae for conditional power. Without this assumption, the hazard ratio within the non-mixture models would be time dependent which causes problems when deriving the asymptotic distribution of the test statistic. Using likelihood methods, it is easy to evaluate appropriateness of the proportional hazard assumption. In our clinical example, this was the case for all assumed parametric families, cf. Section C.1 in the Appendix for the non-mixture exponential model. Due to the high flexibility of the parametric families considered, we believe that the proportional hazard assumption is appropriate for the majority of practical applications.

5.4 Practical Application

We applied our formulae to a data set of a clinical trial for which an interim analysis was performed. We observed considerable differences between power estimates based on a simple exponential model and our non-mixture models. The exponential model clearly did not fit the data well and resulted in an over-optimistic estimate of the conditional power. In contrast, differences between the non-mixture models were small. In general, non-mixture Weibull or non-mixture Gamma models offer a larger flexibility in modelling survival curves compared to the non-mixture exponential model. In practice, we expect larger differences between the exponential,

the non-mixture exponential, and the non-mixture Weibull/Gamma models, while the difference between the non-mixture Weibull and the non-mixture Gamma model might be without practical relevance. We demonstrated that AIC can be used to decide which of the parametric models fits the data best. The analysis plan of clinical trials is usually prescribed in the study protocol. Therefore, we recommend the choice of an appropriate parametric model in the planning phase of a clinical trial. This could be identified e.g. on the basis of available historical data observed under the standard treatment. Here, AIC can be an aid for predetermination of a suitable survival model.

5.5 Aspects of Study Design

Our method allows to quantify the difference in power between the exponential and the non-mixture models. Furthermore, our formulae can be used to calculate sample sizes if the shape of the standard arm is known and a relevant hazard-ratio is specified. For conditional power analysis, we assumed that the time point of final analysis is fixed. However, event-driven designs are frequently used alternatives. These can be addressed with our formulae using the hazard functions estimated at the time point of interim analysis. The time point of final analysis is calculated on the basis of these hazards. Typically, this results in different time points for final analysis for different parametric models. E.g. for our example, final analysis was planned for 14.2 months after interim analysis. Alternatively, assuming an event-driven design, 108 events are required to achieve a power of 80%. Under the non-mixture exponential model, this is achieved at 15.2 months after interim analysis resulting in a conditional power of 19.4% (formerly 18.6%). For the exponential model, the time point of final analysis is 18.0 months after interim analysis resulting in a power of 57.0% (formerly 39.2%). Hence, differences in conditional power between models could even be larger for an event-driven design due to shifts in timing of the final analysis. However, this does not hold in general since the exponential model “generates” less events in the early phase but more events in the later phase compared to the non-mixture exponential model. We like to mention here that event-driven designs have the disadvantage that the required number of events might be unachievable if survival fractions or losses to follow-up are underestimated.

5.6 Non-inferiority Trials

The theory and methods proposed in this paper refer to superiority trials. Another important field of application might be non-inferiority trials. We plan to generalise our methods to this situation in the future. In contrast to superiority trials, a clinically relevant non-inferiority margin needs to be defined prior to sample size calculation. As a consequence, definition of test statistics, hypotheses, and derivations of conditional power formulae are more complex because of a composite null hypothesis. For this situation, the (conditional) power depends on both, the non-inferiority margin Δ and the hazard ratio θ .

5.7 Practical Interpretation of Conditional Power Estimates

When interpreting conditional power, one needs to keep in mind the following: Low conditional power indicates low chance to achieve a significant result at the end of the study. This suggests that termination of the study may be warranted. When the alternative is superior to the standard arm, this would result in high power. However, this does not necessarily imply that the trial should be continued. If the alternative is already significant at time of interim analysis (e.g. applying alpha spending), the trial must be terminated due to ethical reasons.

We like to remark that premature termination of clinical trials due to futility is a complex and serious issue. Conditional power is only one particular indicator which can be used. Confidence intervals of the observed effect sizes – here, the observed hazard ratio – is another measure to decide whether the originally postulated hazard ratio is realistic or if the trial should be stopped for futility. Defining strict boundaries for study termination is difficult for both methods as it depends on time and frequency of interim analyses. Moreover, there are several non-statistical aspects which have to be considered: Premature termination of the study can be decided for the above mentioned ethical reasons or if other treatments become standard. Practical reasons could be an over-estimated recruitment rate or an under-estimated loss to follow-up. Economic aspects could also be relevant, e.g. expensive experimental treatments showing low benefit could more likely be discarded. Stopping for futility is often a mixture of statistical methods and other aspects. In particular, for our clinical example, the major reason for stopping the trial was not the reduced conditional power at interim analysis but the fact that treating patients without the novel drug Rituximab was considered ethically not acceptable. However, we believe that in general conditional power is a useful method to support decision-making in clinical trials.

Acknowledgements

This publication is supported by LIFE – Leipzig Research Center for Civilization Diseases, Leipzig University. This project was funded by means of the European Social Fund and the Free State of Saxony. We thank the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL) directed by Prof. Dr. Michael Pfreundschuh very much for providing data of the highCHOEP trial. We thank our colleague Dr. Peter Ahnert very much for language polishing. We also thank the editor and the reviewers very much for their valuable comments and suggestions to improve the paper.

Conflict of Interest: None declared.

Appendix A: Theorems

Theorem 1

$$\sqrt{n} \frac{W - \log \theta}{\sqrt{2}} \xrightarrow{d} N(0, 1)$$

Proof. Consider the test statistic W according to eq. (1) in Section 2.2 which can be rewritten as

$$W = \log \hat{\theta} = \log \hat{\lambda}_2 - \log \hat{\lambda}_1.$$

The asymptotic normality of the maximum likelihood estimators (Greene [17], p. 478/479) gives

$$\sqrt{n_i} (\hat{\lambda}_i - \lambda_i) \xrightarrow{d} N(0, I^{-1}(\lambda_i))$$

with $I(\lambda_i)$ as the Fisher information. Here, its value can be calculated as

$$I(\lambda_i) = E \left[\frac{d}{d\lambda_i} \log \{ \lambda_i \exp(-\lambda_i T) \} \right]^2 = \lambda_i^{-2}.$$

With the delta method (Serfling [18], p. 122) applied to the logarithm one obtains

$$\sqrt{n_i} (\log \hat{\lambda}_i - \log \lambda_i) \xrightarrow{d} N(0, \lambda_i^{-2} \lambda_i^2) = N(0, 1).$$

With $n = n_1 + n_2$, $W = \log \hat{\lambda}_2 - \log \hat{\lambda}_1$ and $\log \theta = \log \lambda_2 - \log \lambda_1$, this completes the proof.

Theorem 2. It holds that

$$\sqrt{n^*} \frac{W^* - \left(\log \frac{D_2 + \theta \lambda_1 E O_2^*}{O_2 + E O_2^*} - \log \frac{D_1 + \lambda_1 E O_1^*}{O_1 + E O_1^*} \right)}{\sqrt{2}} \xrightarrow{d} N(0, 1),$$

given $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$.

Proof. The proof of Theorem 2 is similar to that of Theorem 1. The main difference here is that one has to consider conditional probabilities, means and variances. Let $t^* > 0$ be the time of interim analysis where data $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$ is available. The test statistic is

$$W^* = \log \frac{D_2 + D_2^*}{O_2 + O_2^*} - \log \frac{D_1 + D_1^*}{O_1 + O_1^*}$$

which is asymptotically normal due to the asymptotic normality of the maximum likelihood estimators (Greene [17], p. 478/479) and the delta method (Serfling [18], p. 122). For the correct centering one has to calculate the (asymptotic) conditional expected value of the test statistic. Therefore, one has

$$\begin{aligned} E \left[\log \frac{D_i + D_i^*}{O_i + O_i^*} \mid (D_i, O_i) = (d_i, o_i) \right] &\approx \log E \left[\frac{D_i + D_i^*}{O_i + O_i^*} \mid (D_i, O_i) = (d_i, o_i) \right] \\ &\approx \log \frac{E [D_i + D_i^* \mid (D_i, O_i) = (d_i, o_i)]}{E [O_i + O_i^* \mid (D_i, O_i) = (d_i, o_i)]} \\ &= \log \frac{d_i + E D_i^*}{o_i + E O_i^*} \end{aligned}$$

by the delta method (Serfling [18], p. 122). Furthermore, with $E\hat{\lambda}_i \rightarrow \lambda_i$ and $\hat{\lambda}_i \approx D_i^*/O_i^*$, one has $\lambda_i \approx E(D_i^*/O_i^*) \approx ED_i^*/EO_i^*$ and so

$$ED_i^* \approx \lambda_i EO_i^*.$$

With $\lambda_2 = \theta\lambda_1$, the proof can be completed.

Theorem 3.

$$\sqrt{n} \frac{W - \log \theta}{\sqrt{\{1 - c_1(1 + \log^2 c_1)\}^{-1} + \{1 - c_1^\theta(1 + (\theta \log c_1)^2)\}^{-1}}} \xrightarrow{d} N(0, 1)$$

Proof. Maximum likelihood estimates for the rate parameter and the cure fraction are

$$\hat{\lambda}_i = \frac{\sum_{j=1}^{n_i} \Delta_{i,j_i}}{\sum_{j=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i} - \log(\hat{c}_i) \sum_{j=1}^{n_i} Y_{i,j_i} \exp(-\hat{\lambda}_i Y_{i,j_i})}$$

and

$$\hat{c}_i = \exp \left[- \frac{\sum_{j=1}^{n_i} \Delta_{i,j_i}}{\sum_{j=1}^{n_i} \{1 - \exp(-\hat{\lambda}_i Y_{i,j_i})\}} \right]$$

respectively (see 2 in Section B of the Appendix).

Due to the same procedure as in the proof of Theorem 1 – here, considering the modified hazard ratio $\theta = \log c_2 / \log c_1$ instead of $\theta = \lambda_2 / \lambda_1$ – the proof of Theorem 3 is omitted. Only the Fisher information must be recalculated:

$$\begin{aligned} I(c_i) &= E \left[\frac{\partial}{\partial c_i} \log \left\{ -c_i^{1-\exp(-\lambda_i T)} \lambda_i \exp(\lambda_i T) \log c_i \right\} \right]^2 \\ &= -\lambda_i \log c_i \left\{ (\lambda_i c_i \log c_i)^{-1} + (\lambda_i c_i \log^3 c_i)^{-1} + (\lambda_i c_i^2 \log^3 c_i)^{-1} \right\}. \end{aligned}$$

Using $I^{-1}(c_i)$ and multiplying it with the square of the first derivative of $\log(-\log)$ (delta method (Serfling [18], p. 122)) gives the (asymptotic) variance of the test statistic. However, one should be cautious when using the asymptotic normality of the maximum likelihood estimators. Since one has only quasi probability density functions within the non-mixture models, it is not obvious that the regularity conditions (Greene [17], p. 474) for the asymptotic properties of the maximum likelihood estimators are fulfilled. In this case they are. The proof of the regularity conditions is beyond the scope of this paper and is available from the corresponding author by request.

Theorem 4. It holds that

$$\sqrt{n^*} \frac{W^* - \left(\log \frac{D_2 - \theta \log(c_1) EO_2^*}{O_2' + EO_2^*} - \log \frac{D_1 - \log(c_1) EO_1^*}{O_1' + EO_1^*} \right)}{\sqrt{\sigma^2}} \xrightarrow{d} N(0, 1),$$

with

$$\begin{aligned} \sigma^2 &= \left\{ 1 - \exp \left(- \frac{D_1 - \log(c_1) EO_1^*}{O_1' + EO_1^*} \right) \left(1 + \left(\frac{D_1 - \log(c_1) EO_1^*}{O_1' + EO_1^*} \right)^2 \right) \right\}^{-1} \\ &\quad + \left\{ 1 - \exp \left(- \frac{D_2 - \theta \log(c_1) EO_2^*}{O_2' + EO_2^*} \right) \left(1 + \left(\frac{D_2 - \theta \log(c_1) EO_2^*}{O_2' + EO_2^*} \right)^2 \right) \right\}^{-1}, \end{aligned}$$

given $(D_1, O_1', D_2, O_2') = (d_1, o_1', d_2, o_2')$.

Proof. The proof of Theorem 4 is omitted because it is similar to the proof of Theorem 3, taking into account Theorem 2. $\square$

Appendix B: Auxiliary results

Lemma 1. Within the non-mixture exponential model with rate parameter $\lambda > 0$ and cure fraction $c \in (0, 1)$, the hazard function $h: [0, \infty) \rightarrow [0, \infty)$ yields

$$h(t) = -\lambda \exp(-\lambda t) \log c.$$

Proof. For deriving the hazard function, the probability density function and the survival function of the non-mixture exponential model are required. For the survival function S it holds $S(t) = c^{1-\exp(-\lambda t)}$. The probability density function can be calculated in the following way:

$$f(t) = -S'(t) = -c^{1-\exp(-\lambda t)} \lambda \exp(-\lambda t) \log c.$$

Division by $S(t)$ yields the supposition. $\square$

Lemma 2 The maximum likelihood estimators for the rate parameter λ and the cure fraction c within the non-mixture exponential model under censoring are

$$\hat{\lambda} = \frac{\sum_{j=1}^n \Delta_j}{\sum_{j=1}^n \Delta_j Y_j - \log(\hat{c}) \sum_{j=1}^n Y_j \exp(-\hat{\lambda} Y_j)}$$

and

$$\hat{c} = \exp \left[-\frac{\sum_{j=1}^n \Delta_j}{\sum_{j=1}^n \{1 - \exp(-\hat{\lambda} Y_j)\}} \right]$$

Proof. In the case of independent observations under consideration of censoring, the likelihood function L is of the form

$$L = \prod_{j=1}^n h^{\delta_j}(t_j) S(t_j),$$

with h as the hazard function and S as the survival function, according to Klein and Moeschberger ([19] p. 74–76). Applying the logarithm one has in the non-mixture exponential case

$$l(\lambda, c) = \sum_{j=1}^n [\delta_j \{\log \lambda - \lambda t_j + \log(-\log c)\} + \{1 - \exp(-\lambda t_j)\} \log c].$$

Considering the score

$$\frac{\partial}{\partial \lambda} l(\lambda, c) = \lambda^{-1} \sum_{j=1}^n \delta_j - \sum_{j=1}^n \delta_j t_j + \log(c) \sum_{j=1}^n t_j \exp(-\lambda t_j)$$

and

$$\frac{\partial}{\partial c} l(\lambda, c) = (c \log c)^{-1} \sum_{j=1}^n \delta_j + c^{-1} \sum_{j=1}^n \{1 - \exp(-\lambda t_j)\}$$

yields the desired formula of the estimators when setting the score to zero and substituting δ_j and t_j by Δ_j respectively Y_j . $\square$

Remark 1. Maximum likelihood estimates within the non-mixture model with Weibull type survival are

$$\begin{aligned}\hat{\lambda}_i &= \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i}^{\hat{k}_i} - \log(\hat{c}_i) \sum_{j_i=1}^{n_i} Y_{i,j_i}^{\hat{k}_i} \exp\left(-\hat{\lambda}_i Y_{i,j_i}^{\hat{k}_i}\right)}, \\ \hat{k}_i &= \left\{ \sum_{j_i=1}^{n_i} \Delta_{i,j_i} \right\} / \left\{ \hat{\lambda}_i \sum_{j_i=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i}^{\hat{k}_i} \log Y_{i,j_i} - \sum_{j_i=1}^{n_i} \Delta_{i,j_i} \log Y_{i,j_i} \right. \\ &\quad \left. - \hat{\lambda}_i \log(\hat{c}_i) \sum_{j_i=1}^{n_i} \exp\left(-\hat{\lambda}_i Y_{i,j_i}^{\hat{k}_i}\right) Y_{i,j_i}^{\hat{k}_i} \log Y_{i,j_i} \right\} \\ \hat{c}_i &= \exp \left[-\frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \left\{ 1 - \exp\left(-\hat{\lambda}_i Y_{i,j_i}^{\hat{k}_i}\right) \right\}} \right].\end{aligned}$$

Remark 2. Applying the maximum likelihood method, the following estimators are derived within the non-mixture model with Gamma type survival:

$$\begin{aligned}\frac{\Gamma(\hat{a}_i)}{\Gamma'(\hat{a}_i)} &= \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\log(\hat{b}_i) \sum_{j_i=1}^{n_i} \Delta_{i,j_i} + \sum_{j_i=1}^{n_i} \Delta_{i,j_i} \log Y_{i,j_i} + \log(\hat{c}_i) \sum_{j_i=1}^{n_i} \frac{\partial}{\partial a_i} \Gamma^0(\hat{a}_i, \hat{b}_i Y_{i,j_i})}, \\ \hat{b}_i &= \frac{\hat{a}_i \sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i} - \log(\hat{c}_i) \sum_{j_i=1}^{n_i} \frac{\partial}{\partial b_i} \Gamma^0(\hat{a}_i, \hat{b}_i Y_{i,j_i})} \\ \hat{c}_i &= \exp \left\{ -\frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \Gamma^0(\hat{a}_i, \hat{b}_i Y_{i,j_i})} \right\}.\end{aligned}$$

Remark 3. When calculating the hazard functions h_{Wei} and h_{Gam} for the non-mixture Weibull and the non-mixture Gamma model in the same way as it is done in the proof of Lemma 1, one obtains

$$\begin{aligned}h_{\text{Wei}}(t) &= -\lambda k t^{k-1} \exp(-\lambda t^k) \log c \\ h_{\text{Gam}}(t) &= -\frac{b^a}{\Gamma(a)} t^{a-1} \exp(-bt) \log c.\end{aligned}$$

A Non-mixture models

A.1 Accuracy of proportional hazard assumption

For visualisation of differences in estimating the survival curves with and without assuming proportional hazards, we provide Figure 6 in case of the non-mixture exponential model.

Figure 6: Comparison of fitted survival curves of the non-mixture exponential model. Left: Parameters are estimated without proportional hazard assumption (AIC = 702.9311). Right: Parameters are estimated with proportional hazard assumption $\lambda_1 = \lambda_2$ (AIC = 702.9322).

Appendix C.2: The Weibull-type survival: conditional power formula

We consider the hypotheses $H_0: \theta = 1$ vs. $H_1: \theta \neq 1$ together with eq. (4). With an analogue procedure as in Section 3, assuming $\lambda_1 = \lambda_2$ and $k_1 = k_2$, i.e. the proportional hazard assumption, the asymptotic conditional power function is of the form eq. (3):

$$\gamma_c(\theta) = \Phi\left(\frac{q_{\frac{\alpha}{2}}\sigma_{H_0} + \mu_{H_0} - \mu_{H_1}(\theta)}{\sigma_{H_1}(\theta)}\right) + \left\{1 - \Phi\left(\frac{q_{1-\frac{\alpha}{2}}\sigma_{H_0} + \mu_{H_0} - \mu_{H_1}(\theta)}{\sigma_{H_1}(\theta)}\right)\right\}$$

with $O_i^* = \sum_{j=1}^{n_i^*} \left\{1 - \exp\left(-\hat{\lambda}_i Y_{i,j_i}^* \hat{k}_i\right)\right\}$ as the corresponding substitute within the parameters of the conditional power function defined at the end of Section 3.3:

$$\begin{aligned} n\mu_{H_0} &= \log\left(\frac{d_2 - \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right) - \log\left(\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right), \\ \sigma_{H_0}^2 n &= \left[n^* \left\{1 - \exp\left(-\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right) \left(1 + \left(\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right)^2\right)\right\}\right]^{-1} \\ &\quad + \left[n^* \left\{1 - \exp\left(-\frac{d_2 - \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right) \left(1 + \left(\frac{d_2 - \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right)^2\right)\right\}\right]^{-1}, \\ \mu_{H_1}(\theta) &= \log\left(\frac{d_2 - \theta \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right) - \log\left(\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right) \end{aligned}$$

and

$$\begin{aligned} \sigma_{H_1}^2(\theta) &= \left[n^* \left\{1 - \exp\left(-\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right) \left(1 + \left(\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right)^2\right)\right\}\right]^{-1} \\ &\quad + \left[n^* \left\{1 - \exp\left(-\frac{d_2 - \theta \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right) \left(1 + \left(\frac{d_2 - \theta \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right)^2\right)\right\}\right]^{-1}. \end{aligned}$$

Appendix C.3: The Gamma-type survival: conditional power formula

The hypotheses are as usual, $H_0: \theta = 1$ vs. $H_1: \theta \neq 1$, together with eq. (4). Again, similar to Section 3, assuming proportional hazards, i.e. $a_1 = a_2$ and $b_1 = b_2$, one obtains the asymptotic conditional power function (3):

$$\gamma_c(\theta) = \Phi\left(\frac{q_{\frac{\alpha}{2}}\sigma_{H_0} + \mu_{H_0} - \mu_{H_1}(\theta)}{\sigma_{H_1}(\theta)}\right) + \left\{1 - \Phi\left(\frac{q_{1-\frac{\alpha}{2}}\sigma_{H_0} + \mu_{H_0} - \mu_{H_1}(\theta)}{\sigma_{H_1}(\theta)}\right)\right\}$$

with $O_i^* = \sum_{j_i=1}^{n_i} \Gamma^0(\hat{a}_i, \hat{b}_i Y_{i,j_i}^*)$ as the substitution within the parameters of the conditional power function presented in Section 3.3:

$$\begin{aligned} * \mu_{H_0} &= \log\left(\frac{d_2 - \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right) - \log\left(\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right), \\ \sigma_{H_0}^2 &= \left[n^* \left\{ 1 - \exp\left(-\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right) \left(1 + \left(\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right)^2 \right) \right\} \right]^{-1} \\ &\quad + \left[n^* \left\{ 1 - \exp\left(-\frac{d_2 - \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right) \left(1 + \left(\frac{d_2 - \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right)^2 \right) \right\} \right]^{-1}, \\ \mu_{H_1}(\theta) &= \log\left(\frac{d_2 - \theta \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right) - \log\left(\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right) \end{aligned}$$

and

$$\begin{aligned} \sigma_{H_1}^2(\theta) &= \left[n^* \left\{ 1 - \exp\left(-\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right) \left(1 + \left(\frac{d_1 - \log(c_1)EO_1'^*}{o_1' + EO_1'^*}\right)^2 \right) \right\} \right]^{-1} \\ &\quad + \left[n^* \left\{ 1 - \exp\left(-\frac{d_2 - \theta \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right) \left(1 + \left(\frac{d_2 - \theta \log(c_1)EO_2'^*}{o_2' + EO_2'^*}\right)^2 \right) \right\} \right]^{-1}. \end{aligned}$$

References

- [1] Lambert PC. Modeling of the cure fraction in survival studies. *Stata J.* 2007;7:351–375.
- [2] Andersen PK. Conditional power calculations as an aid in the decision whether to continue a clinical trial. *Control Clin Trials.* 1987;8:67–74.
- [3] Martinez EZ, Achcar JA, Jácome AA, Santos JS. Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application to gastric cancer data. *Compute Methods Programs Biomed.* 2013;112:343–355.
- [4] Tsodikov AD, Ibrahim JG, Yakovlev AY. Estimating cure rates from survival data: an alternative to two-component mixture models. *J Am Stat Assoc.* 2003;98:1063–1078.
- [5] Tsodikov A, Loeffler M, Yakovlev A. A cure model with time-changing risk factor: an application to the analysis of secondary Leukaemia. *Stat Med.* 1998;17:27–40.
- [6] Akaike H. A new look at the statistical model identification. *IEEE Trans Automatic Control.* 1974;19:716–723.
- [7] Pfreundschuh M, Zwick C, Zeynalova S, Duehrsen U, Pflueger KH, Vrieling T, et al. Dose-escalated CHOEP for the treatment of young patients with aggressive non-Hodgkin's lymphoma: II. Results of the randomized high-CHOEP trial of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). *Ann Oncol.* 2007;19:545–552.
- [8] Altman DG, Bland JM. Parametric v non-parametric methods for data analysis. *BMJ.* 2009;338:a3167.
- [9] Achcar JA, Coelho-Barros EA, Mazucheli J. Cure fraction models using mixture and non-mixture models. *Tatra Mountains Math Publications.* 2012;51:1–9.
- [10] Boag JW. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J R Stat Soc Ser B Method.* 1949;11:15–53.
- [11] Berkson J, Gage RP. Survival curve for cancer patients following treatment. *J Am Stat Assoc.* 1952;47:501–515.
- [12] Sposto R, Sather HN. Determining the duration of comparative clinical trials while allowing for cure. *J Chronic Dis.* 1985;38:683–690.
- [13] Sposto R. Cure model analysis in cancer: an application to data from the children's cancer group. *Stat Med.* 2002;21:293–312.
- [14] Ibrahim JG, Chen MH, Sinha D. Bayesian semiparametric models for survival data with a cure fraction. *Biometrics.* 2001;57:383–388.
- [15] Yin G, Ibrahim JG. A general class of bayesian survival models with zero and nonzero cure fractions. *Biometrics.* 2005;61:403–412.
- [16] Cox DR. Regression models and life-tables. *J R Stat Soc Ser B Methodol.* 1972;34:187–220.
- [17] Greene WH. Econometric analysis. Upper Saddle River: Prentice Hall, 2003.
- [18] Serfling RJ. Approximation theorems of mathematical statistics. New York: John Wiley & Sons, 1980.
- [19] Klein JP, Moeschberger ML. Survival analysis. Techniques for censored and truncated data. New York: Springer Science+Business Media, 2005.

Supplemental Material: The online version of this article (DOI:ijb-2015-0073) offers supplementary material, available to authorized users.