NEXRES 9 (2026) 101713

Contents lists available at ScienceDirect

Next Research

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

Research Paper

Time-dependent hazard ratios in breast cancer clinical trials

Roman Tomaschitz

Sechsschimmelgasse 1/21-22, A-1090 Vienna, Austria

ARTICLE INFO

Keywords:
Multiparametric survival functions
Cumulative hazard functions and varying hazard ratios
Confidence bands and time-dependent *p*-values
Nonlinear multiparameter regression
Kaplan−Meier data sets

ABSTRACT

Multiply broken power-law densities are employed as cumulative hazard functions for the analytic modeling of Kaplan−Meier (KM) data sets from oncology randomized clinical trials. These adaptive densities are applied to KM data sets for disease-free survival (DFS) from the IBCSG 22-00 trial, which evaluated maintenance therapy with cyclophosphamide and methotrexate for HER2-negative early breast cancer. Also studied are KM data sets for overall survival (OS) and progression-free survival (PFS) from the CLEOPATRA trial, which examined combination therapy with pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer. The multiparameter survival functions of the experimental and control trial arms are regressed and compared with Weibull fits. The uniform local accuracy of the regressed distributions over the entire empirical time range determined by the uncensored event times is demonstrated by residual plots. Confidence bands of the survival and cumulative hazard functions are calculated as well as 95% confidence intervals for selected percentiles. The time evolution of the cumulative-hazard ratio of the trial arms is compared with the constant hazard ratio from the semiparametric Cox regression model.

1. Introduction

The purpose of this paper is to introduce adaptive survival functions that can accurately model Kaplan−Meier (KM) data sets from randomized clinical trials in oncology, cf., e.g., Refs. [1–3], covering the entire data range and allowing extrapolation beyond the uncensored event times. The principal aim is the precise analytic modeling of the hazard ratio of the experimental and control trial arms, in particular the time evolution of this ratio, its confidence band and varying p-value. This will be done without resorting to linearization around a baseline hazard, in contrast to the Cox proportional hazards model and its generalizations, cf., e.g., Refs. [4–7].

The limitations of linearization in analyzing clinical research data have recently been discussed in Refs. [8,9]. The strengths and drawbacks of generalized Cox regression with fractional polynomials or restricted cubic splines have been studied in the review [10]. These models admit time-dependent hazard ratios while maintaining the linear dependence of the log-hazard on the fitting parameters. Here, we will abandon the Cox regression ansatz and use nonlinear regression from the outset. The hazard functions of the trial arms are not any more generated by adding small perturbations to a baseline and can be largely differing functions; they are elementary closed-form expressions with nonlinear parameter dependence admitting analytic extrapolation outside the data range.

The cumulative hazards defining the survival functions $S(t) = \mathrm{e}^{-H(t)}$ are modeled as multiply broken power-law densities, $H(t) = a_0 t^{\alpha_0} \prod_{k=1}^{n} \left(1 + (t/b_k)^{\beta_k/|\eta_k|}\right)^{\eta_k}$, with positive parameters a_0, b_k, β_k, real α_0 and non-zero exponents η_k inferred from least-squares regression. The number n of factors in the product can be freely chosen. As pointed out below, these densities have polygonal chain limits which are flexible enough to approximate complex data sets extending over several orders in time.

The uncensored and censored event times of the experimental and control arms of the breast cancer trials analyzed here are taken from Ref. [1] and converted to KM data points and Greenwood variances, cf., e.g., Refs. [4–7], defining the weight factors of the least-squares functional of the survival function. Specifically, we will use KM data for disease-free survival (DFS) from the IBCSG 22-00 trial [11] investigating cyclophosphamide and methotrexate (CM) maintenance chemotherapy for human epidermal growth factor receptor 2 (HER2)-negative early breast cancer. As a second example, we will consider KM data for overall survival (OS) and progression-free survival (PFS) from the CLEOPATRA trial [12–14] testing the efficacy of adding pertuzumab to trastuzumab and docetaxel therapy for HER2-positive metastatic breast cancer.

Multiply broken power-law distributions defining the cumulative hazard functions are particularly effective to model the staircase-like data structures of PFS studies emerging in Log−Log plots. In the limit

E-mail address: tom@geminga.org.

https://doi.org/10.1016/j.nexres.2026.101713
Received 9 November 2025; Received in revised form 17 March 2026; Accepted 2 April 2026
Available online 3 April 2026

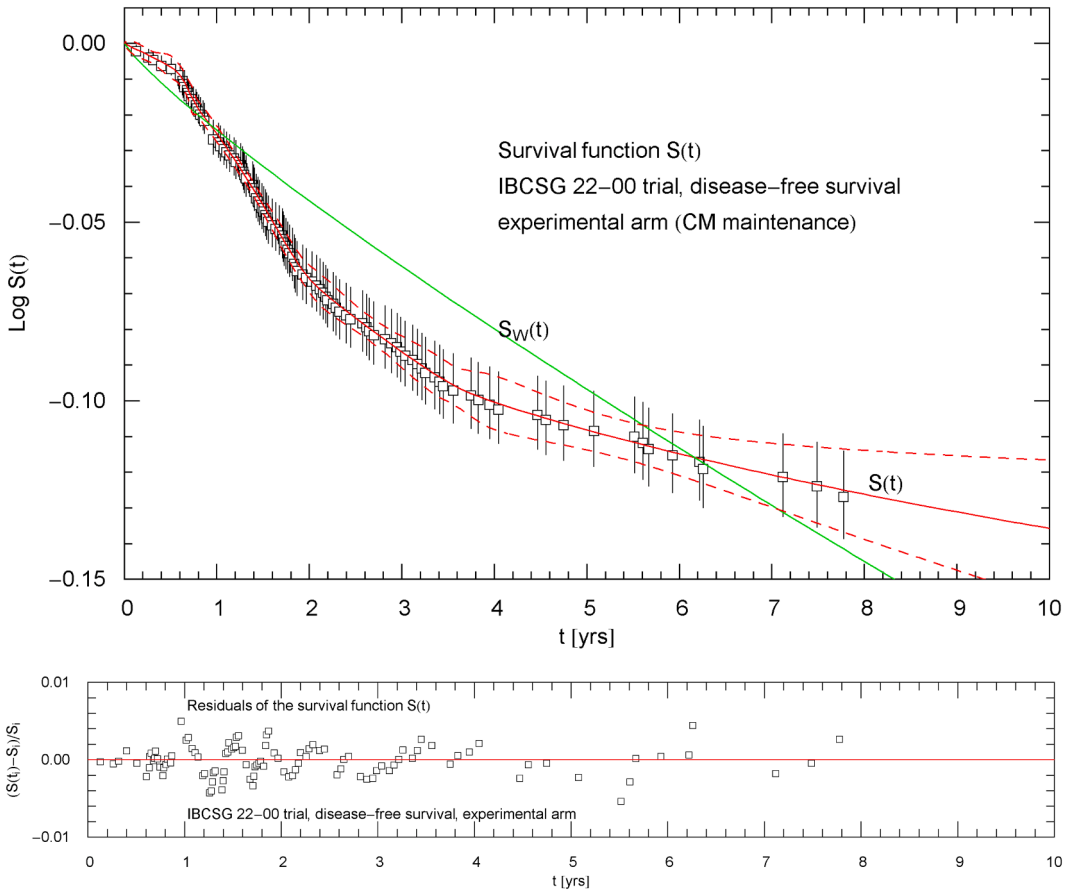

Fig. 1. Survival function $S(t)$ for DFS of the experimental arm (CM maintenance) of the IBCSG 22–00 trial [11]. KM data points from Ref. [1]. The open squares depict the Kaplan−Meier (KM) data points of the survival function $S(t)$ in Log−Lin coordinates, cf. Section 2. The error bars show the Greenwood standard deviations of the KM points. The red solid curve depicts the regressed analytic survival function $S(t)$, which is an exponentiated broken power-law distribution, with fitting parameters and goodness-of-fit parameters in Table 1. The red dashed curves are the boundaries of the 95% (1.96 σ) confidence band of $S(t)$. The Weibull fit $S_W(t)$ of Ref. [2] (green solid curve, cf. Section 2) is also indicated for comparison. The lower panel shows the residuals of the regressed $S(t)$ with respect to the data points. This residual plot is more informative than the goodness-of-fit parameters in Table 1, as it quantifies the local accuracy of the fit.

of vanishing exponents $\eta_k \to 0$, the density $H(t)$ degenerates into a polygonal chain (polyline) when plotted in double-logarithmic coordinates, the parameters b_k (break points) of $H(t)$ defining the location of the vertices. Conversely, for every polygonal chain defining a single-valued function, there exists a density $H(t)$ which admits this chain as limit. The parameters of this density, including the number of factors, are uniquely defined by the prescribed polygonal chain, except the η_k exponents, which approach zero in this limit. This property can be used to find initial values for the iterative minimization of the nonlinear least-squares functional, just by approximating the data set initially by a polygonal chain, which is then smoothed out by the least-squares iteration. Since a data set defining an empirical curve without excessive scatter can usually be approximated by a polygonal chain, the above stated distributions have wide applicability beyond the modeling of survival data.

In Section 2, the least-squares functional (multiparametric, nonlinear and weighted) is defined and the analytic survival and cumulative hazard functions of the mentioned trials are regressed, together with their 95% confidence bands and residual plots.

In Section 3, selected survival percentiles and survival rates of the experimental and control arms of the trials are calculated, including confidence intervals. This is done for DFS studied in the IBCSG 22–00 trial and for OS and PFS in the CLEOPATRA trial.

In Section 4, the time evolution of the cumulative-hazard ratio of the trial arms is discussed, the DFS ratio of the IBCSG 22–00 trial and the OS and PFS ratios of the CLEOPATRA trial. The 95% confidence bands of the ratios are calculated, and the varying ratios are compared with the constant ratios from the Cox proportional hazards model. We present plots depicting the time evolution of the ratios and also record the ratios at selected survival times, their confidence intervals and p-values.

Section 5 contains the conclusions. In Appendix A, polygonal chain limits of multiply broken power laws are discussed, which supply initial values for the nonlinear multiparameter regression of the survival functions. An overview and summary of this paper is also provided by the captions of Figs. 1–15 and Tables 1–6.

2. Multiparametric analysis of the IBCSG 22-00 and CLEOPATRA trial data: survival and cumulative hazard functions

The KM data points $(t_i, S_i)_{i=1,\ldots,N}$ of the survival function $S(t)$ of the experimental and control arm of the IBCSG 22–00 trial [11] (with DFS as primary end point) are depicted in Figs. 1 and 3. In these figures, Log−Lin coordinates are used, where simple exponentials $ae^{\alpha t}$ ($a > 0$, real α) appear as straight lines. The error bars $(t_i, S_i \pm \sigma_i)_{i=1,\ldots,N}$, on the data points depict the standard deviations σ_i defined by the Greenwood variances of the KM points. We used electronic data files from Refs. [1] listing uncensored and censored event times, which were obtained by digitalizing KM plots [15–17] in the trial papers [11,12]. The KM data points of the survival functions of the CLEOPATRA trial arms are depicted in Figs. 6 and 8 for OS and Figs. 11 and 13 for PFS.

The data points $(t_i, H_i)_{i=1,\ldots,N}$, $H_i = -\log S_i$, of the cumulative hazard function $H(t)$ of the experimental and control arm of the IBCSG 22–00

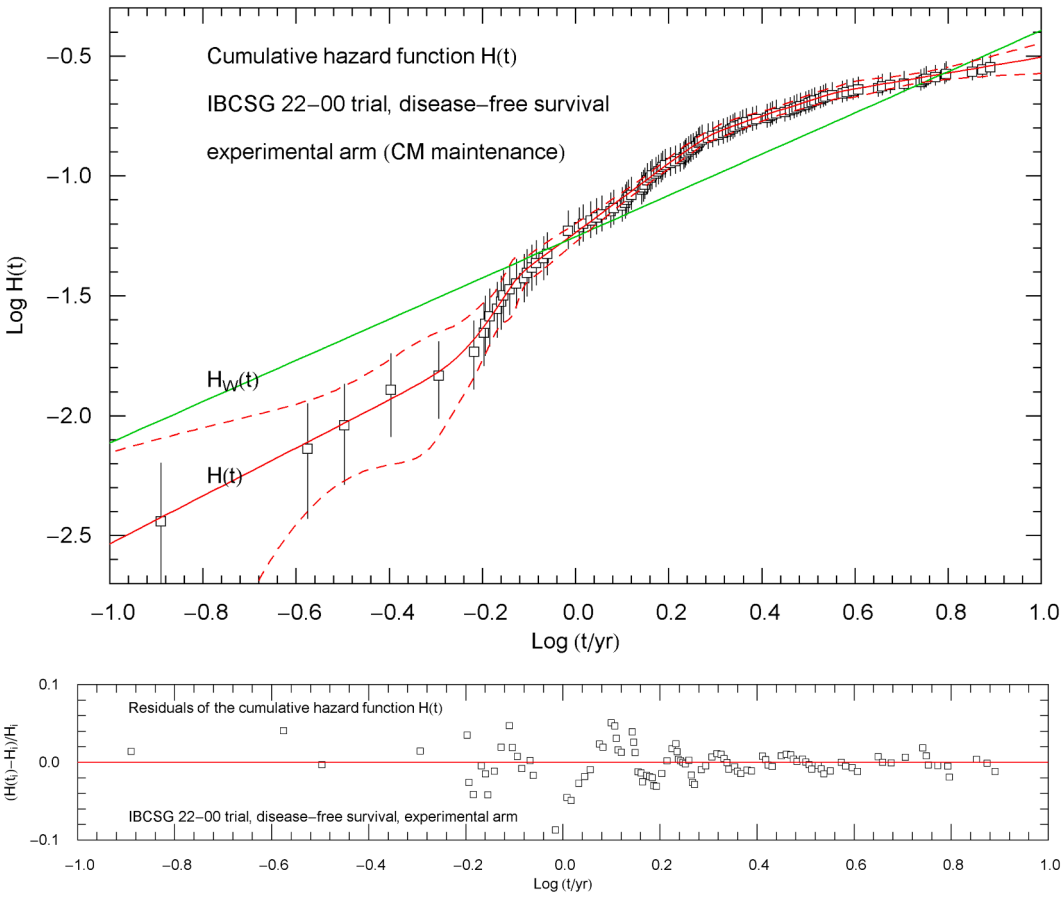

Fig. 2. Cumulative hazard function $H(t)$ for DFS of the experimental arm (CM maintenance) of the IBCSG 22–00 trial [11]. The data points are obtained from the KM data of the survival function, and the error bars from the Greenwood standard deviations, cf. Section 2. The cumulative hazard function (red solid curve) is the broken power-law distribution $H(t) = -\log S(t)$ in Eq. (2.1) and Table 1, derived from the regressed survival function $S(t)$, cf. Fig. 1. The red dashed curves are the boundaries of the 95% confidence band of $H(t)$. The Weibull cumulative hazard function $H_{\rm W}(t)$ (green line), obtained from the Weibull survival function $S_{\rm W}(t)$, see Fig. 1, is also depicted for comparison, cf. Ref. [2]. $H_{\rm W}(t)$ is a simple power law, appearing as straight line in the Log–Log coordinates of this figure. The lower panel shows the residuals of the regressed $H(t)$ curve with respect to the data points.

trial are plotted in Figs. 2 and 4. Log–Log coordinates are employed in this case, where simple power laws at^α ($a > 0$, real α) appear as straight lines. The standard deviations $\widehat{\sigma}_i = \sigma_i/S_i$ of the $(t_i, H_i)_{i=1,\ldots,N}$ points define the error bars $(t_i, H_i \pm \widehat{\sigma}_i)_{i=1,\ldots,N}$ indicated in the figures. The data points of the cumulative hazard functions of the CLEOPATRA trial arms are depicted in Figs. 7 and 9 for OS and Figs. 12 and 14 for PFS.

The hazard functions of the experimental and control trial arms are modeled as multiply broken power-law distributions (cf. Refs. [18–20]),

$$H(t) = a_0 t^{\alpha_0} \prod_{k=1}^{n} \left(1 + (t/10^{c_k})^{\beta_k/|\eta_k|}\right)^{\eta_k}, \tag{2.1}$$

with positive amplitude a_0, real exponents α_0, c_k, non-zero exponents η_k, and positive exponents β_k. The number n of factors in the product can be adapted to the data set. The regression of the survival functions $S(t) = \exp(-H(t))$ is performed with the least-squares functional

$$\chi^2_S\Big(a_0, \alpha_0, (c_k, \beta_k, \eta_k)_{k=1,\ldots,n}\Big) = \sum_{i=1}^{N} \frac{(S(t_i) - S_i)^2}{\sigma_i^2}. \tag{2.2}$$

To find initial values for the iteration of this functional, we consider the $\eta_k \to 0$ limit of density (2.1), which consist of power-law segments appearing as polygonal chain when plotted in decadic Log–Log coordinates, cf. Appendix A. The initial parameters are reconstructed from the polygonal chain approximating the empirical curve defined by the data points.

We briefly sketch how this polygonal chain limit emerges. Density $H(t)$ in (2.1) is composed of factors $p_k(t) = \left(1 + (t/b_k)^{\beta_k/|\eta_k|}\right)^{\eta_k}$, $b_k := 10^{c_k}$. The $\eta_k \to 0$ limit of these factors reads $p_k(t) \sim [t < b_k] + [b_k \le t](t/b_k)^{\beta_k s_k}$, with $s_k = \mathrm{sign}(\eta_k)$. Here, we use Iverson brackets, which evaluate to one if the inequality/equality in them is satisfied and to zero otherwise. The $\eta_k \to 0$ limit of density (2.1) can thus be assembled as

$$H(t) \sim \sum_{k=0}^{n} [b_k \le t < b_{k+1}] a_0 t^{\alpha_0} \prod_{i=1}^{k} (t/b_i)^{\beta_i s_i}, \tag{2.3}$$

with $b_0 = 0$ and $b_{n+1} = +\infty$. This limit evidently consists of power-law segments. Therefore, in Log–Log plots, the $\eta_k \to 0$ limit of $H(t)$ is a polygonal chain. There are no constraints on this chain other than defining a one-valued function, cf. Appendix A.

The regressed survival functions $S(t)$ of the IBCSG 22–00 trial are shown in Figs. 1 and 3 (for DFS, experimental and control arm, red solid curves) and the survival functions of the CLEOPATRA trial arms are depicted in Figs. 6 and 8 (for OS) and Figs. 11 and 13 (for PFS). The cumulative hazard functions $H(t)$ of the IBCSG 22–00 trial arms (for DFS) are shown in Figs. 2 and 4. The hazard functions of the CLEOPATRA trial arms (for OS and PFS) are depicted in Figs. 7,9 and Figs. 12,14, respectively.

A Mathematica® [21] routine to minimize χ^2_S in (2.2) is FindMinimum [chisquared [...],{initial values},MaxIterations→nmax]. The regressed parameters a_0, α_0, $(c_k, \beta_k, \eta_k)_{k=1,\ldots,n}$ of $S(t)$ and $H(t)$ are

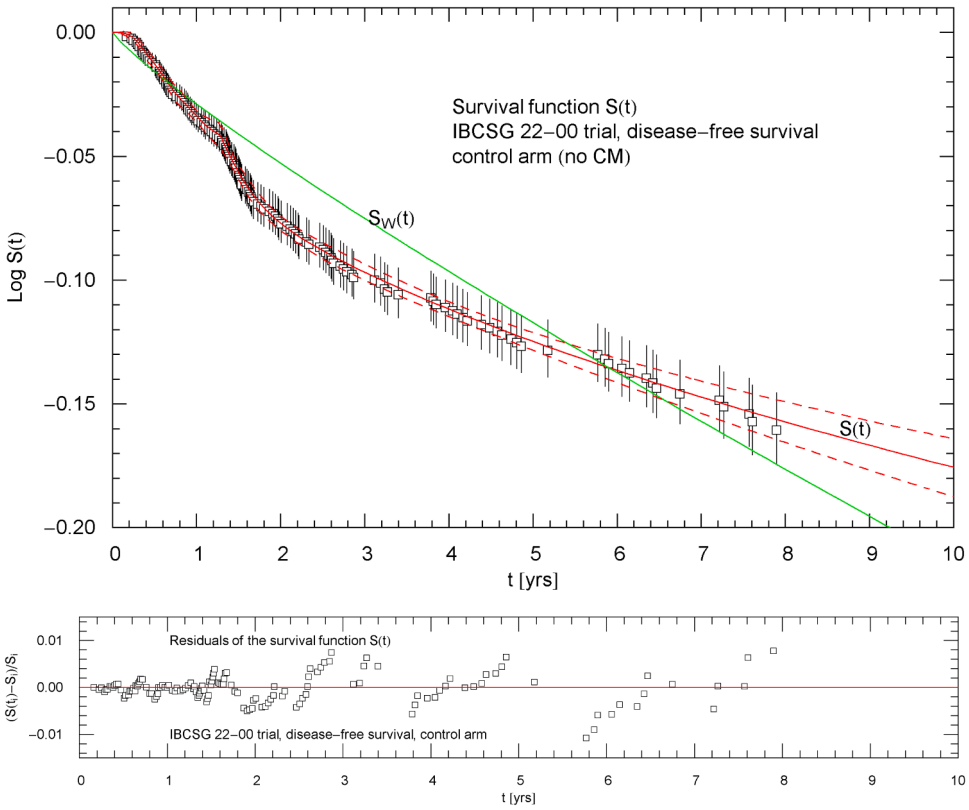

Fig. 3. Survival function $S(t)$ for DFS of the control arm (no CM maintenance) of the IBCSG 22–00 trial [11]. KM data points from Ref. [1]. The error bars show the Greenwood standard deviations. The red solid curve depicts the regressed analytic survival function $S(t)$, cf. Table 1. The red dashed curves are the boundaries of the 95% confidence band of $S(t)$. For comparison, $S_{\rm W}(t)$ (green solid curve) is a Weibull fit from Ref. [2]. Residuals are plotted in the lower panel.

recorded in Table 1 for the enumerated trial arms. The coefficient of determination R^2, cf. Ref. [22], the standard error SE, the minimum of the least-squares functional χ^2_S, and the degrees of freedom (dof) are also listed in Table 1. Also shown in Figs. 1–4,6–9,11–14 are the 95% confidence bands of $S(t)$ and $H(t)$ (red dashed boundary curves), calculated with the covariance matrix obtained from the weighted least-squares functional (2.2), cf., e.g., Ref. [23]. Residual panels quantifying the local deviation of the regressed analytic distributions $S(t)$ and $H(t)$ from the data points are included in the figures.

For comparison, we also reproduced the Weibull fits of Ref. [2] (green solid curves in Figs. 1–4,6–9,11–14). The Weibull survival function $S_{\rm W}(t)=\exp\big(-\big(t/{\rm e}^{\beta_{\rm w}}\big)^{\alpha_{\rm w}}\big)$ of the IBCSG 22–00 trial, with parameters $\alpha_{\rm w}=0.86$, $\beta_{\rm w}=5.84$ for the DFS experimental arm and $\alpha_{\rm w}=0.87$, $\beta_{\rm w}=5.60$ for the control arm, is depicted in Figs. 1 and 3. The Weibull cumulative hazard function is a simple power law, $H_{\rm W}(t)={\rm e}^{-\alpha_{\rm w}\beta_{\rm w}}t^{\alpha_{\rm w}}$ (with the above parameters for experimental and control arm), which appears as straight line in the Log−Log coordinates of Figs. 2 and 4, cf., e.g., Ref. [24]. The unit of time in these Weibull functions is 1 month, whereas the time unit in Figs. 1-4 is 1 year, which requires multiplying the time variable by a factor of 12 when plotting these functions.

The Weibull parameters of the CLEOPATRA trial are $\alpha_{\rm w}=1.30$, $\beta_{\rm w}=4.35$ for the OS experimental arm and $\alpha_{\rm w}=1.32$, $\beta_{\rm w}=4.04$ for the OS control arm. Regarding progression-free survival, the Weibull parameters of this trial are $\alpha_{\rm w}=1.10$, $\beta_{\rm w}=3.48$ for the PFS experimental arm and $\alpha_{\rm w}=1.07$, $\beta_{\rm w}=3.09$ for the PFS control arm, cf. Ref. [2].

3. Survival percentiles and survival rates

Table 2 lists selected percentiles $t_{k/100}$ of the regressed survival functions of the experimental and control arms of the IBCSG 22–00 trial, for DFS. The percentiles denote the time at which disease recurrence is detected in k% of patients in the respective trial arm. The $t_{k/100}$ and their confidence intervals (CIs) are calculated from the regressed survival functions, cf. the caption of Table 2, and are within the time range depicted in Figs. 1,3. Since the survival functions are analytic and in closed form, it is possible to extrapolate beyond the data range as done in the figures. However, extrapolation beyond the 10-year range is not meaningful, since the confidence bands of the survival and hazard functions tend to become rather wide outside the data range (see Figs. 1–4) and, therefore, also the confidence intervals of the extrapolated percentiles. (As for the blanks in Table 2, we indicated only percentiles with reasonably narrow confidence intervals.)

Table 4 lists selected percentiles $t_{k/100}$ of the CLEOPATRA trial arms, for OS and PFS. The entry $t_{k/100}$ refers to the time at which death (in the case of OS) or disease progression (in the case of PFS) is observed in k% of patients. Moderate extrapolation beyond the data range in Figs. 6,8,11,13 is also attempted in Table 4, by way of the regressed analytic survival functions.

Median OS in the experimental arm is 57.6 months, with 95% CI [55.3,59.8], cf. Table 4. In the control arm, median OS is 42.4 months, with 95% CI [41.7,43.0]. For comparison, using the Cox regression model, median OS of the experimental arm is quoted in Ref. [12] as 56.5 months and the lower boundary of the 95% CI as 49.3 months. Median OS of the control arm based on the Cox model is 40.8 months with 95% CI [35.8,48.3], cf. Ref. [12].

Table 3 (IBCSG 22-00) and Tables 5 and 6 (CLEOPATRA) list the survival rates (that is, the survival function in percent units rather than normalized to one) of the experimental and control arm of the respective trial at selected survival times, including 95% CIs, which are read off from the confidence bands depicted in Figs. 1,3,6,8,11,13. The survival rates refer to DFS (IBCSG 22-00, Table 3) and OS and PFS (CLEOPATRA, Tables 5 and 6, respectively).

For instance, in the case of the IBCSG 22-00 trial, the 5-year DFS rate

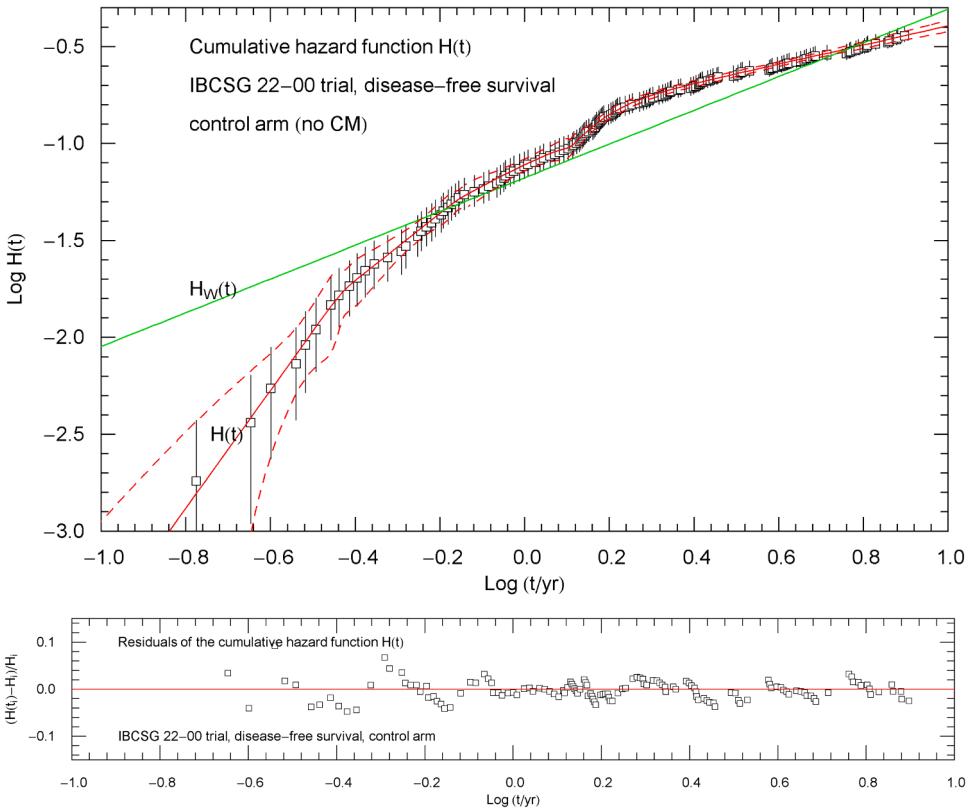

Fig. 4. Cumulative hazard function $H(t)$ for DFS of the control arm (no CM maintenance) of the IBCSG 22–00 trial. The data points are obtained from the KM data of the survival function, and the error bars from the Greenwood standard deviations, cf. Section 2. The cumulative hazard function $H(t)$ (red solid curve, cf. Eq. (2.1)) is derived from the regressed survival function $S(t)$, cf. Fig. 3. The red dashed curves are the boundaries of the 95% confidence band of $H(t)$. The Weibull cumulative hazard function $H_W(t)$ (green solid straight line), obtained from the Weibull survival function $S_W(t)$, see Fig. 3, is also depicted for comparison, cf. Ref. [2]. Residuals of the least-squares fit are plotted in the lower panel.

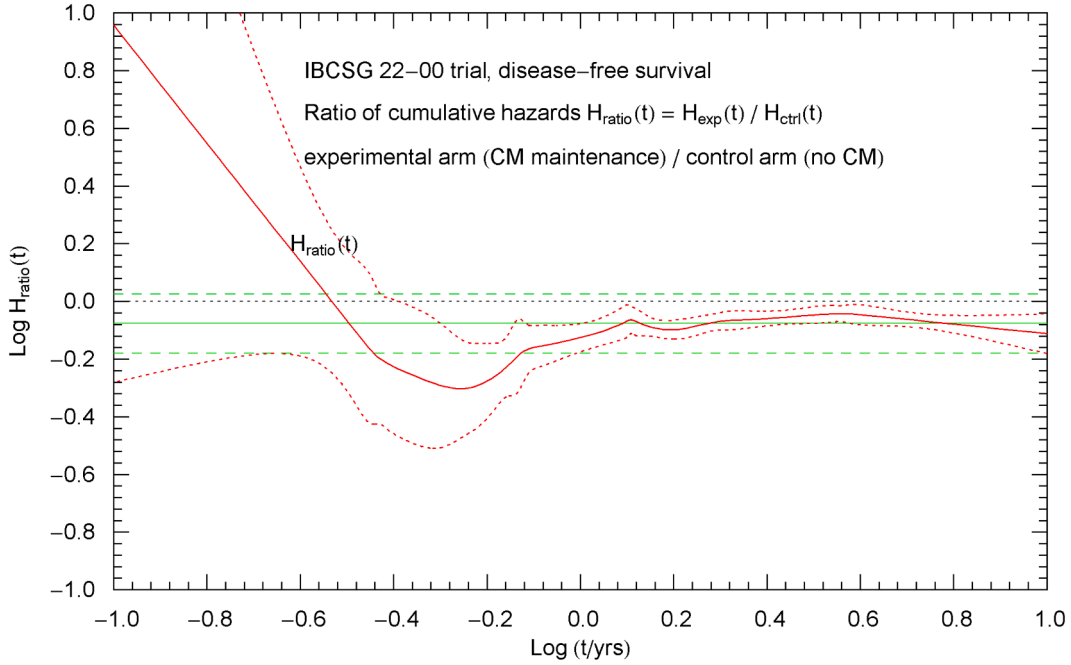

Fig. 5. Time evolution of the cumulative-hazard ratio of the IBCSG 22–00 trial [11], for DFS. Depicted is the ratio $H_{exp}(t)/H_{ctrl}(t)$ (red solid curve, cf. Section 4) of the cumulative hazard functions of the experimental arm $H_{exp}(t)$ and control arm $H_{ctrl}(t)$, cf. Figs. 2,4. The red dotted curves are the boundaries of the 95% confidence band of the ratio. For comparison, the green solid horizontal line shows the constant hazard ratio of the Cox proportional hazards model, and the green dashed horizontal lines depict the 95% confidence interval thereof, cf. Ref. [11].

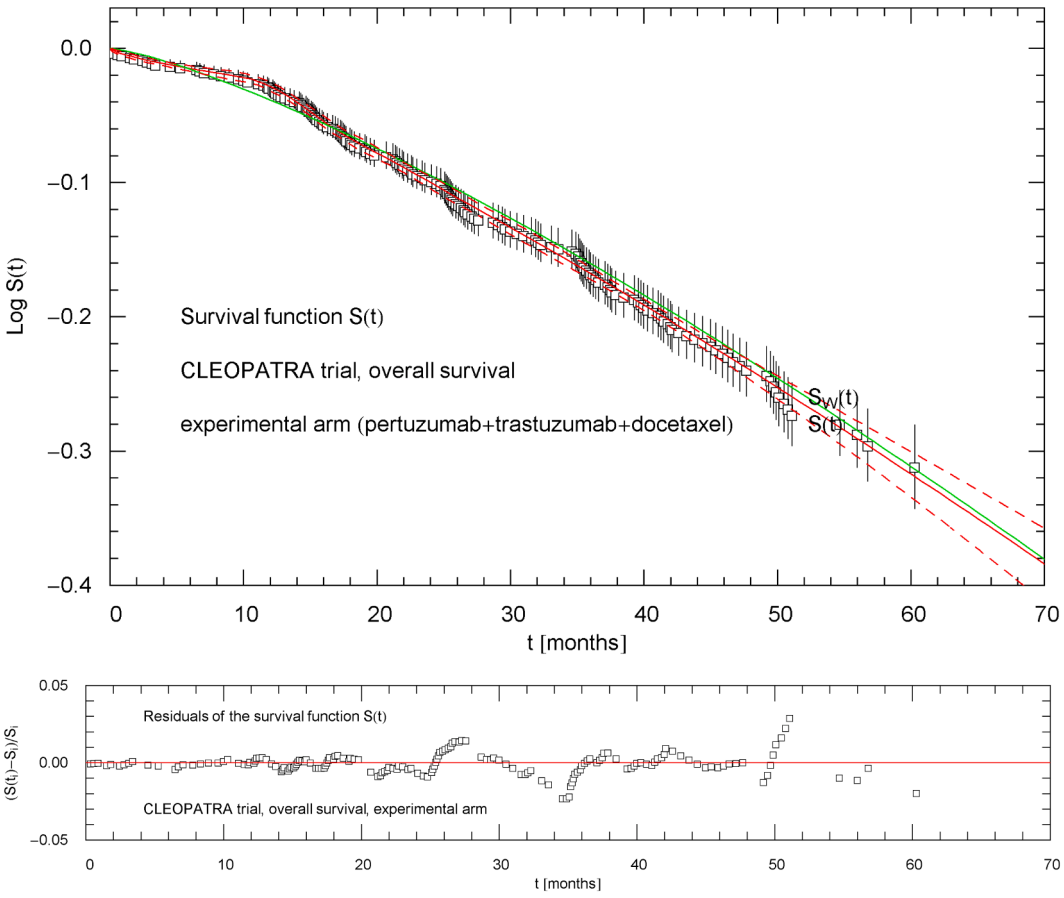

Fig. 6. Survival function $S(t)$ for OS of the experimental arm (pertuzumab + trastuzumab + docetaxel) of the CLEOPATRA trial [12]. KM data points from Ref. [1]. The error bars show the Greenwood standard deviations. The red solid curve depicts the regressed analytic survival function $S(t)$, cf. Section 2. The red dashed curves are the boundaries of the 95% confidence band of $S(t)$. The Weibull fit $S_W(t)$ of Ref. [2] (green solid curve) is also indicated for comparison. Residuals of the least-squares fit are depicted in the lower panel.

of the experimental arm is 78% with 95% CI [77.0,79.0]. The 5-year DFS rate of the control arm is 75.0% with 95% CI [74.4,75.7], obtained from the regressed survival functions. The 5-year DFS rate obtained from the Cox regression model is 78.1% for the experimental arm and 74.7% for the control arm, as quoted in Ref. [11].

4. Time-dependent cumulative-hazard ratios

The cumulative hazard functions $H_{\rm exp}(t)$ and $H_{\rm ctrl}(t)$ of the experimental and control arm of the IBCSG 22-00 trial quantify the risk of disease recurrence, accumulated from the time of randomization. In the case of the CLEOPATRA trial, $H_{\rm exp}(t)$ and $H_{\rm ctrl}(t)$ quantify the accumulated risk of death or disease progression. The ratio $H_{\rm ratio}(t) = H_{\rm exp}(t)/H_{\rm ctrl}(t)$ of the IBCSG 22-00 trial is plotted in Fig. 5 (red solid curve). The 95% confidence band of the ratio is also indicated in this figure by red dotted boundary curves. The time evolution of the cumulative-hazard ratio of the CLEOPATRA trial is depicted in Fig. 10 for OS and Fig. 15 for PFS, including confidence bands.

For comparison, the constant hazard ratio of the Cox proportional hazards model is depicted in Figs. 5,10,15 as green solid horizontal line, and the 95% CI thereof is indicated by the green dashed horizontal lines. The constant Cox hazard ratio for DFS of the IBCSG 22-00 trial in Fig. 5 is 0.84, with 95% CI [0.66,1.06] and p-value of 0.14, as quoted in Ref. [11].

The constant Cox hazard ratio of the CLEOPATRA trial in Fig. 10 (green solid horizontal line) is estimated as 0.68 for OS in Ref. [12], and the green dashed horizontal lines in Fig. 10 show the 95% CI thereof, which is [0.56,0.84] with $p < 0.001$, based on the assumption of proportional hazards. For PFS, the Cox hazard ratio of the CLEOPATRA trial depicted in Fig. 15 (by the green solid horizontal line) is also estimated as 0.68 in Ref. [12], and the horizontal green dashed lines in Fig. 15 indicate the 95% CI thereof, which is [0.58,0.80] with $p < 0.001$, cf. Ref. [12]. The pitfalls of assuming constant hazard ratios are discussed in Refs. [25–29], and p-values are critiqued in Refs. [30–34].

Table 3 (IBCSG 22-00) and Tables 5 and 6 (CLEOPATRA) list the time-dependent cumulative-hazard ratio $H_{\rm ratio}(t) = H_{\rm exp}(t)/H_{\rm ctrl}(t)$ (depicted in Figs. 5,10,15) of the experimental and control arm at selected survival times and the respective 95% CIs as well as the p-value of $\log H_{\rm ratio}(t)$, which is calculated from the z-score as $p(t) = \operatorname{erfc}\big(|z(t)|/\sqrt{2}\big)$ (erfc: complementary error function). If $|z(t)| > 1.96$, the p-value falls below the $\alpha = 0.05$ significance level. The time-dependent z-score reads $|z(t)| = |\log H_{\rm ratio}(t)|/\sigma_{H,\log{\rm ratio}}(t)$, where $\sigma_{H,\log{\rm ratio}}(t)$ is the standard error of $\log H_{\rm ratio}(t)$ defining the boundary curves $H_{\rm ratio}(t)\exp\big[\pm 1.96\sigma_{H,\log{\rm ratio}}(t)\big]$ of the 95% confidence band of $H_{\rm ratio}(t)$, see Figs. 5,10,15. The IBCSG 22-00 ratios in Table 3 refer to DFS, and the CLEOPATRA ratios to OS in Table 5 and to PFS in Table 6.

The upper boundaries of the confidence intervals in Table 3 (which lists the DFS ratio of the IBCSG 22-00 trial) do not exceed one, although only barely so (see also Fig. 5), so that all recorded p-values are below the 0.05 significance level. This is also the case for the CLEOPATRA trial; the upper boundaries of the 95% CIs of the varying OS and PFS hazard ratios listed in Tables 5 and 6 are below one (except the 6-month entry in Table 5) and the corresponding p-values below 0.05. The time-dependent hazard ratios and their confidence intervals in Tables 3,5,6 are to be compared with the constant Cox hazard ratios quoted above.

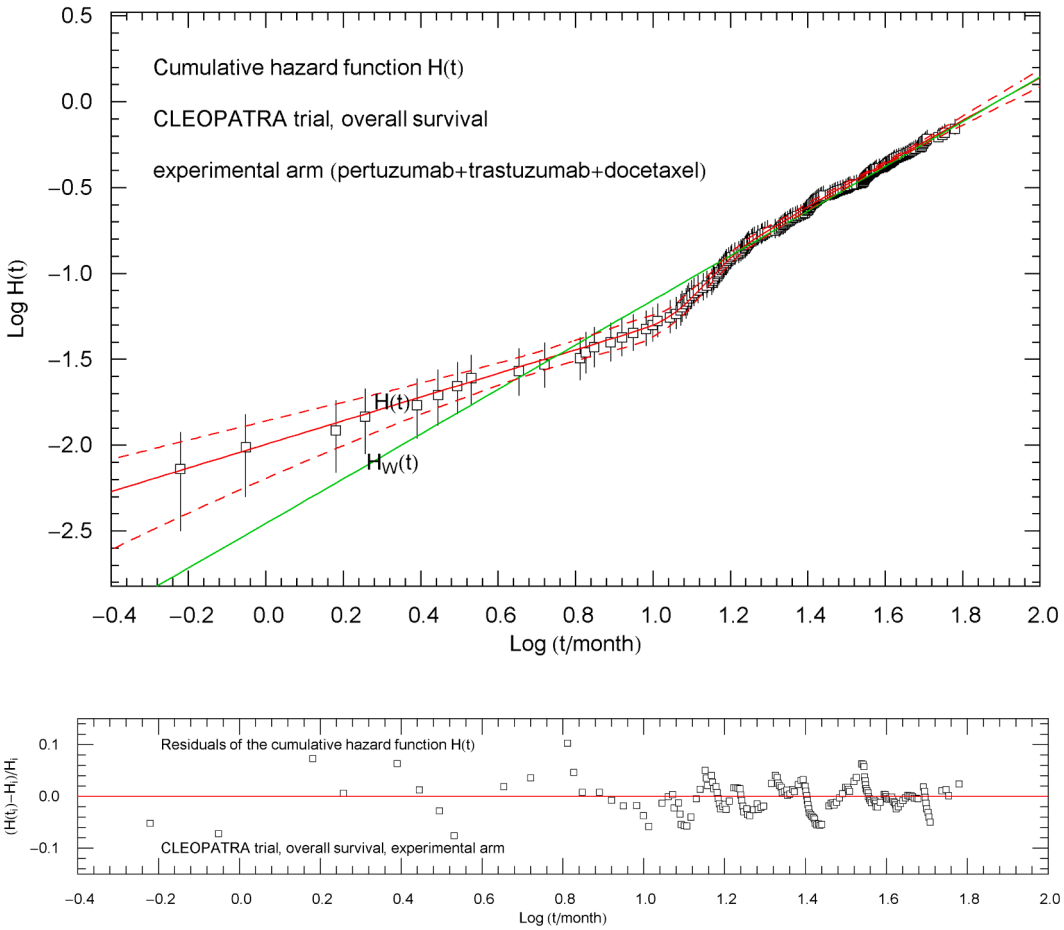

Fig. 7. Cumulative hazard function $H(t)$ for OS of the experimental arm (pertuzumab + trastuzumab + docetaxel) of the CLEOPATRA trial [12]. The data points are obtained from the KM data of the survival function in Fig. 6, and the error bars from the Greenwood standard deviations, cf. Section 2. The cumulative hazard function (red solid curve) is the broken power-law distribution $H(t)$ in Eq. (2.1) and Table 1, derived from the regressed survival function $S(t)$. The red dashed curves are the boundaries of the 95% confidence band of $H(t)$. The Weibull cumulative hazard function $H_W(t)$ (green straight line), obtained from the Weibull survival function $S_W(t)$, see Fig. 6, is also depicted for comparison, cf. Ref. [2]. The lower panel shows the residuals of the least-squares fit.

5. Conclusion

We used adaptive survival functions defined by multiply broken power-law densities for the continuum modeling of KM data sets. The distributions were put to the test with KM data for disease-free survival (DFS) from the IBCSG 22-00 trial [11] and for overall and progression-free survival (OS and PFS) from the CLEOPATRA trial [12]. The multiparametric survival functions of the experimental and control arms of the trials were regressed and their 95% (1.96 σ) confidence bands calculated, see Figs. 1 and 3 (IBCSG 22-00, DFS), Figs. 6 and 8 (CLEOPATRA, OS) and Figs. 11 and 13 (CLEOPATRA, PFS). The survival functions were found to give uniformly accurate analytic representations of the KM data, as demonstrated by the residual panels included in the figures.

Selected percentiles of the survival functions of the trial arms were recorded in Table 2 (IBCSG 22-00, DFS) and Table 4 (CLEOPATRA, OS and PFS). Extrapolation of the analytic survival functions beyond the data range (depicted in Figs. 1,3,6,8,11,13) is only in limits possible, because of the rapid widening of the confidence bands outside the data range. Survival rates for the experimental and control trial arms are listed at selected times in Table 3 (IBCSG 22-00, DFS), Table 4 (CLEOPATRA, OS) and Table 5 (CLEOPATRA, PFS). The confidence intervals of the survival rates of the trial arms do not overlap at the listed survival times, suggesting a statistically significant increase in the survival rates of the experimental DFS/PFS/OS arm.

Figs. 2 and 4 (IBCSG 22-00, DFS) and Figs. 7 and 9 (CLEOPATRA, OS) and Figs. 12,14 (CLEOPATRA, PFS) show the cumulative hazard function $H(t)$ of the trial arms and the corresponding data points as well as residual panels. Fig. 5 (IBCSG 22-00, DFS) and Figs. 10 and 15 (CLEOPATRA, OS and PFS) depict the time variation of the cumulative-hazard ratio $H_{\rm exp}(t)/H_{\rm ctrl}(t)$ of the experimental and control arm and the 95% confidence band of the ratio. In intervals where the upper boundary of the confidence band lies below the time axis (in the Log−Log coordinates of Figs. 5,10,15), the accumulated risk in the experimental arm is significantly reduced at the 95% level as compared with the control arm, and the time-dependent p-value falls below the 0.05 threshold.

The constant ratio inferred from the proportional hazards assumption of the Cox model, cf. Refs. [4,5], is also shown in Figs. 5,10,15 for comparison. Assuming proportionality, the ratios of cumulative and differential hazards coincide, which is not the case for varying hazard ratios. We focused here on ratios of cumulative hazards rather than ratios of differential hazards $H'_{\rm exp}(t)/H'_{\rm ctrl}(t)$, as the accumulated risk (from the time of randomization onward) is more relevant in clinical trials than the risk (of disease recurrence, disease progression, death, etc.) at a specific instant. In Table 3 (IBCSG 22-00, DFS) and Tables 4 and 5 (CLEOPATRA, OS and PFS), the cumulative-hazard ratio $H_{\rm exp}(t)/H_{\rm ctrl}(t)$ is listed at selected survival times, together with the respective 95% confidence interval and p-value. Accurate survival functions with narrow confidence bands are essential for the calculation of time-dependent hazard ratios, since the confidence band of the ratio will be wider due to error propagation, as exemplified in the figures.

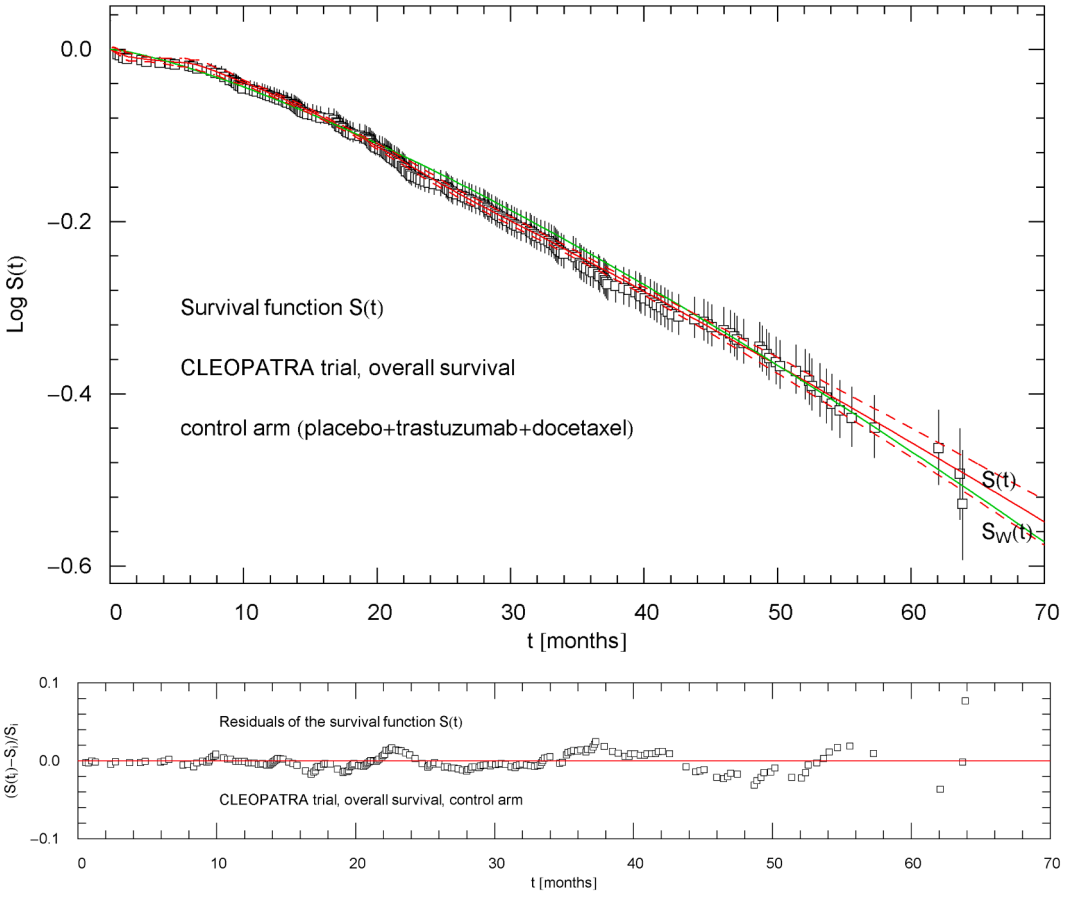

Fig. 8. Survival function $S(t)$ for OS of the control arm (placebo + trastuzumab + docetaxel) of the CLEOPATRA trial [12]. KM data points from Ref. [1]. The error bars show the Greenwood standard deviations. The red solid curve depicts the regressed analytic survival function $S(t)$, cf. Section 2 and Table 1. The red dashed curves are the boundaries of the 95% confidence band of $S(t)$. For comparison, $S_{\rm W}(t)$ (green solid curve) is a Weibull fit from Ref. [2]. Residuals are plotted in the lower panel.

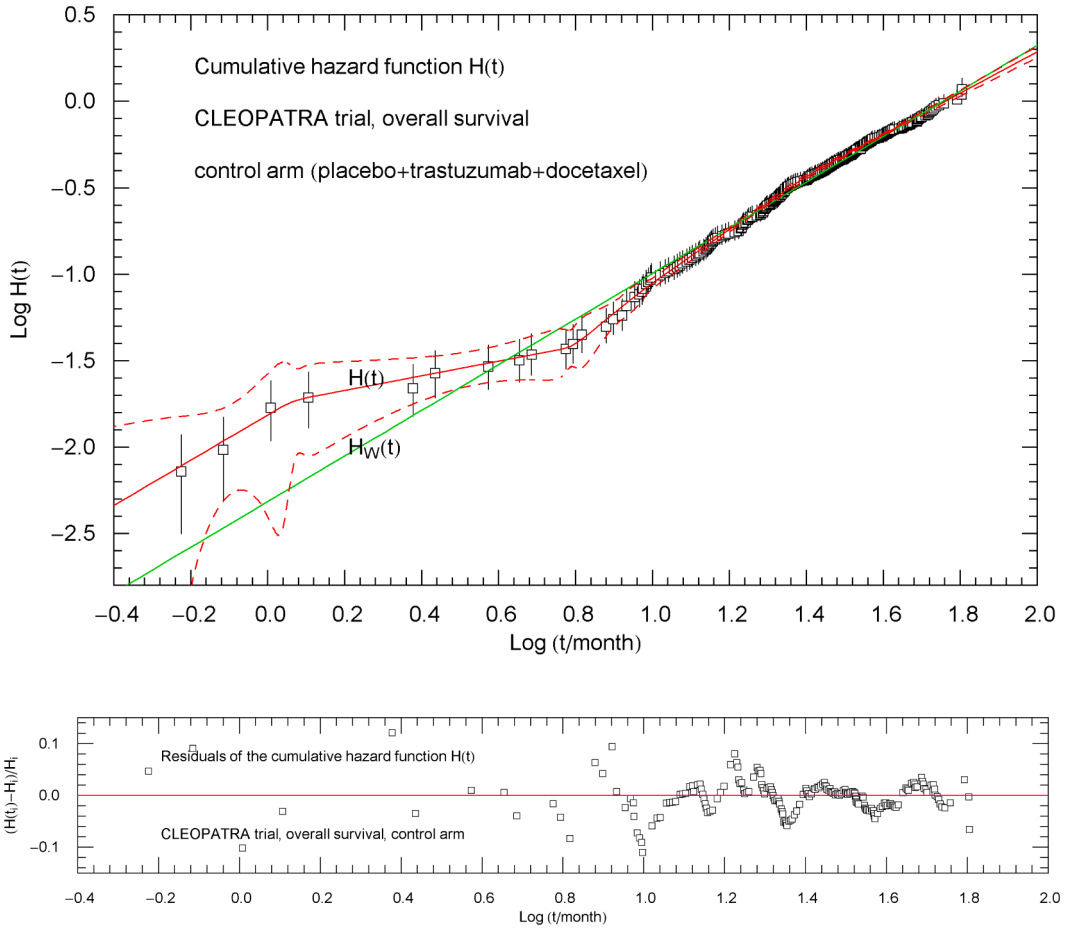

Fig. 9. Cumulative hazard function $H(t)$ for OS of the control arm (placebo + trastuzumab + docetaxel) of the CLEOPATRA trial [12]. The data points are obtained from the KM data of the survival function, and the error bars from the Greenwood standard deviations, cf. Section 2. The cumulative hazard function $H(t)$ (red solid curve) is derived from the regressed survival function $S(t)$ in Fig. 8. The red dashed curves are the boundaries of the 95% confidence band of $H(t)$. The Weibull cumulative hazard function $H_W(t)$ (green straight line), obtained from the Weibull survival function $S_W(t)$ in Fig. 8, is also depicted for comparison, cf. Ref. [2]. Residuals of the least-squares fit are plotted in the lower panel.

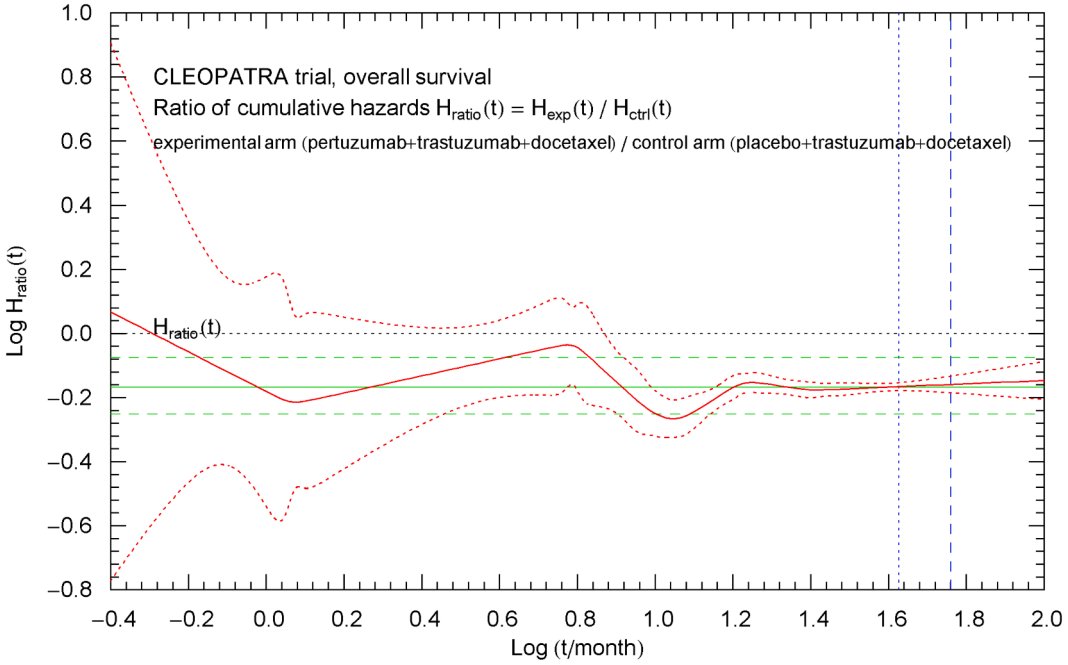

Fig. 10. Time evolution of the cumulative-hazard ratio of the CLEOPATRA trial [12], for OS. Depicted is the ratio $H_{\rm exp}(t)/H_{\rm ctrl}(t)$ (red solid curve, cf. Section 4) of the cumulative hazard functions of the experimental arm $H_{\rm exp}(t)$ and control arm $H_{\rm ctrl}(t)$, cf. Figs. 7,9. The red dotted curves are the boundaries of the 95% confidence band of the ratio. For comparison, the green solid horizontal line shows the constant hazard ratio of the Cox regression model, and the green dashed horizontal lines depict the 95% confidence interval thereof, cf. Ref. [12]. The blue dotted and dashed vertical lines indicate the median survival time of the control and experimental arm, respectively, cf. Table 4.

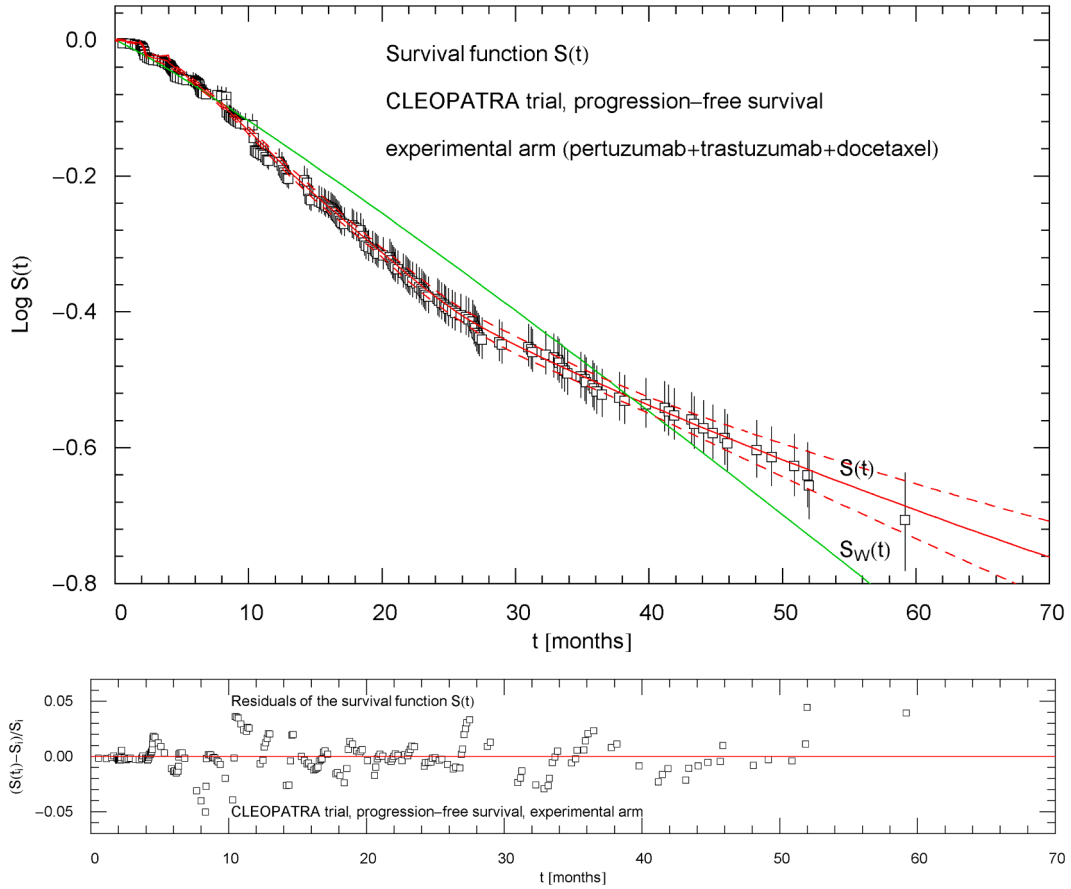

Fig. 11. Survival function $S(t)$ for PFS of the experimental arm (pertuzumab + trastuzumab + docetaxel) of the CLEOPATRA trial [12]. KM data points from Ref. [1]. The error bars show the Greenwood standard deviations. The red solid curve depicts the regressed analytic survival function $S(t)$, cf. Section 2 and Table 1. The red dashed curves are the boundaries of the 95% confidence band of $S(t)$. The Weibull fit $S_{\rm W}(t)$ of Ref. [2] (green solid curve) is also indicated for comparison. Residuals of the least-squares fit are depicted in the lower panel.

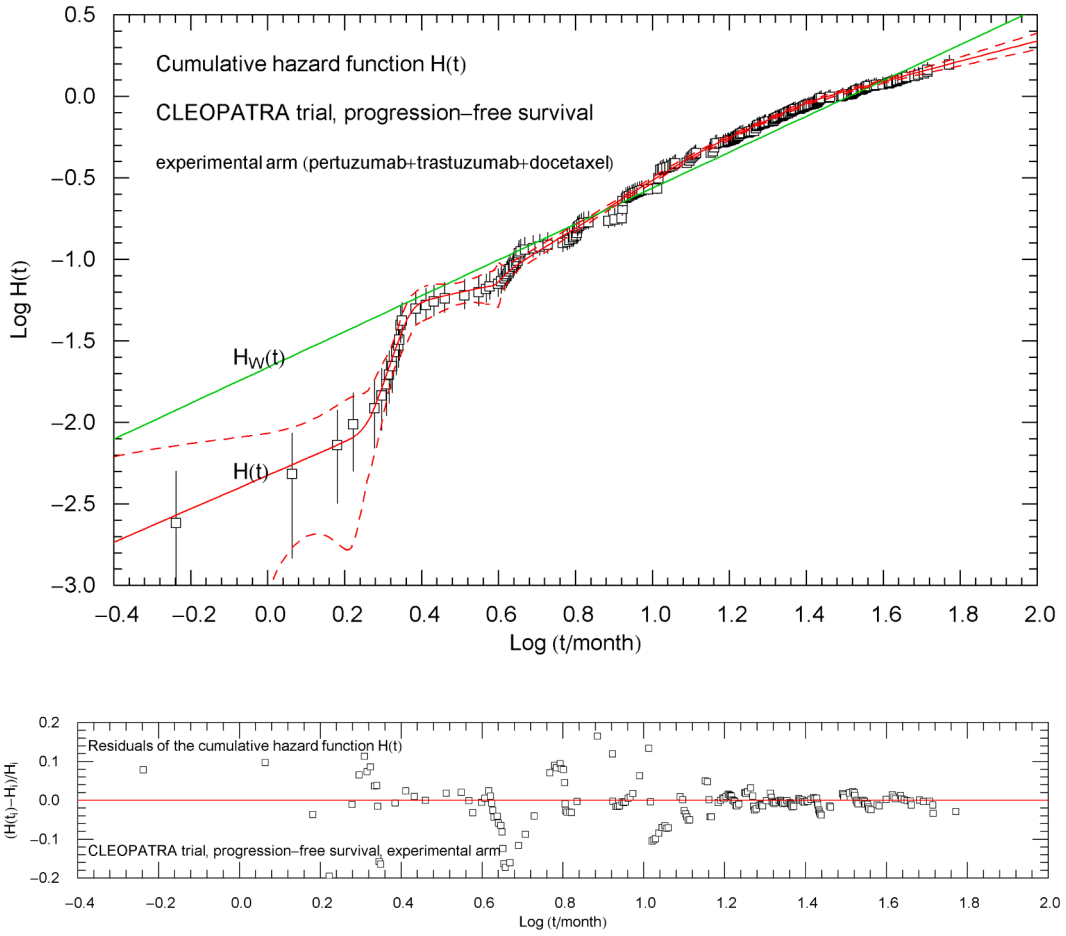

Fig. 12. Cumulative hazard function $H(t)$ for PFS of the experimental arm (pertuzumab + trastuzumab + docetaxel) of the CLEOPATRA trial [12]. The data points are obtained from the KM data of the survival function in Fig. 11, and the error bars from the Greenwood standard deviations, cf. Section 2. The cumulative hazard function (red solid curve) is the broken power-law distribution $H(t)$ in Eq. (2.1) and Table 1, derived from the regressed survival function $S(t)$. The red dashed curves are the boundaries of the 95% confidence band of $H(t)$. The Weibull cumulative hazard function $H_{\rm W}(t)$ (green straight line), obtained from the Weibull survival function $S_{\rm W}(t)$, cf. Fig. 11, is also depicted for comparison, cf. Ref. [2]. The lower panel shows the residuals of the least-squares fit.

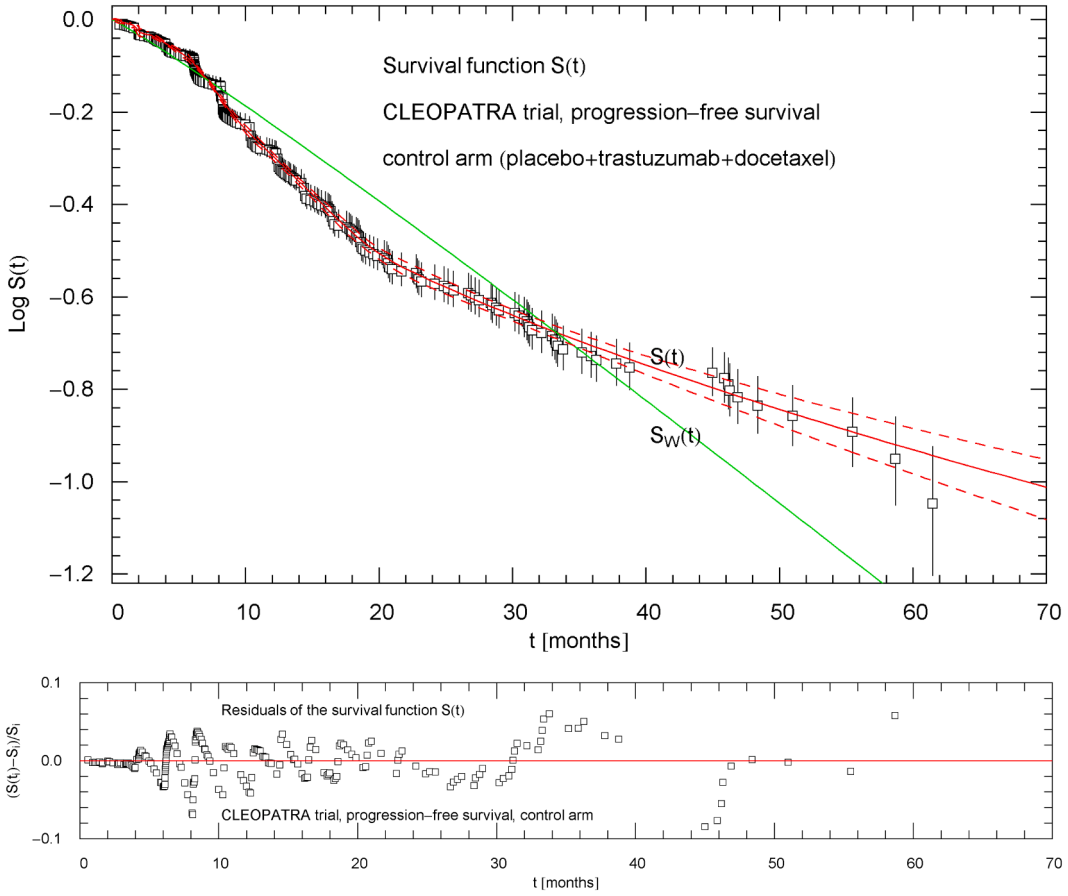

Fig. 13. Survival function $S(t)$ for PFS of the control arm (placebo + trastuzumab + docetaxel) of the CLEOPATRA trial [12]. KM data points from Ref. [1]. The error bars show the Greenwood standard deviations. The red solid curve depicts the regressed analytic survival function $S(t)$, cf. Section 2 and Table 1. The red dashed curves are the boundaries of the 95% confidence band of $S(t)$. For comparison, $S_W(t)$ (green solid curve) is a Weibull fit from Ref. [2]. Residuals are plotted in the lower panel.

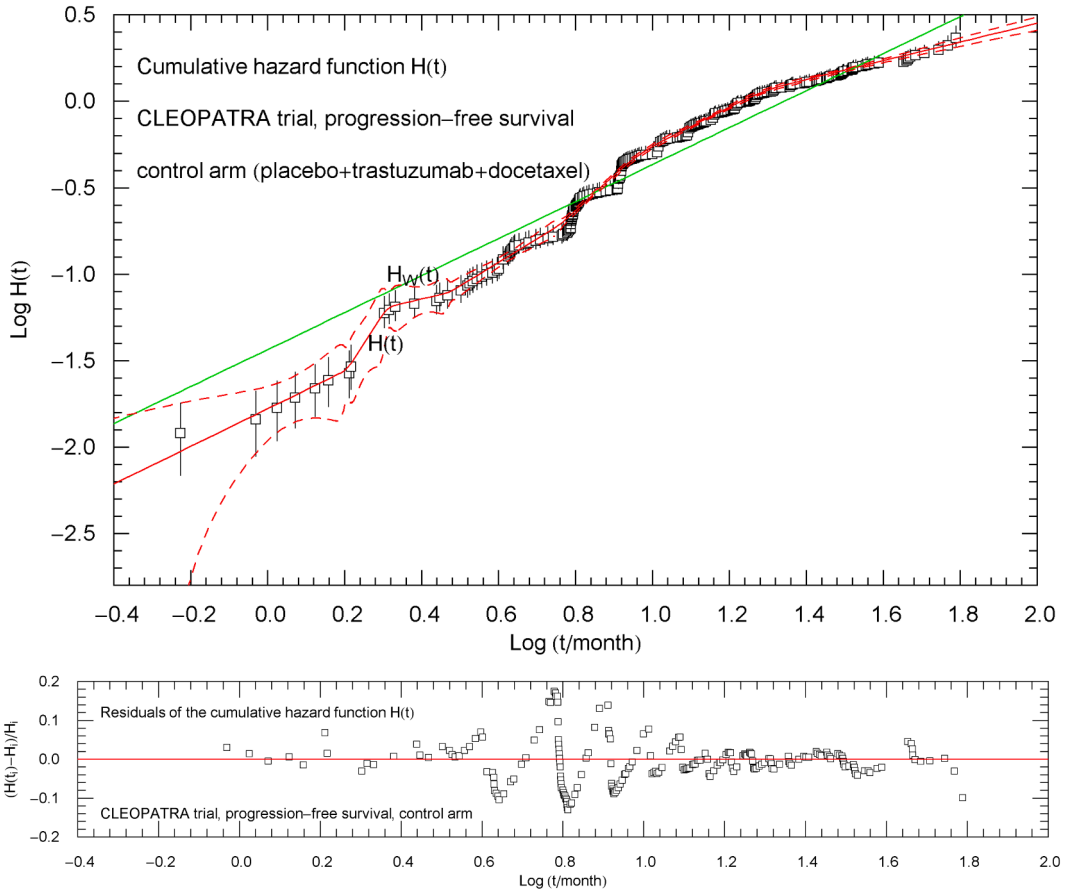

Fig. 14. Cumulative hazard function $H(t)$ for PFS of the control arm (placebo + trastuzumab + docetaxel) of the CLEOPATRA trial [12]. The data points are obtained from the KM data of the survival function, and the error bars from the Greenwood standard deviations, cf. Section 2. The cumulative hazard function $H(t)$ (red solid curve, cf. Eq. (2.1)) is derived from the regressed survival function $S(t)$ in Fig. 13. The red dashed curves are the boundaries of the 95% confidence band of $H(t)$. The Weibull cumulative hazard function $H_W(t)$ (green straight line, cf. Ref. [2]), obtained from the Weibull survival function $S_W(t)$ in Fig. 13, is also depicted for comparison. Residuals of the least-squares fit are plotted in the lower panel.

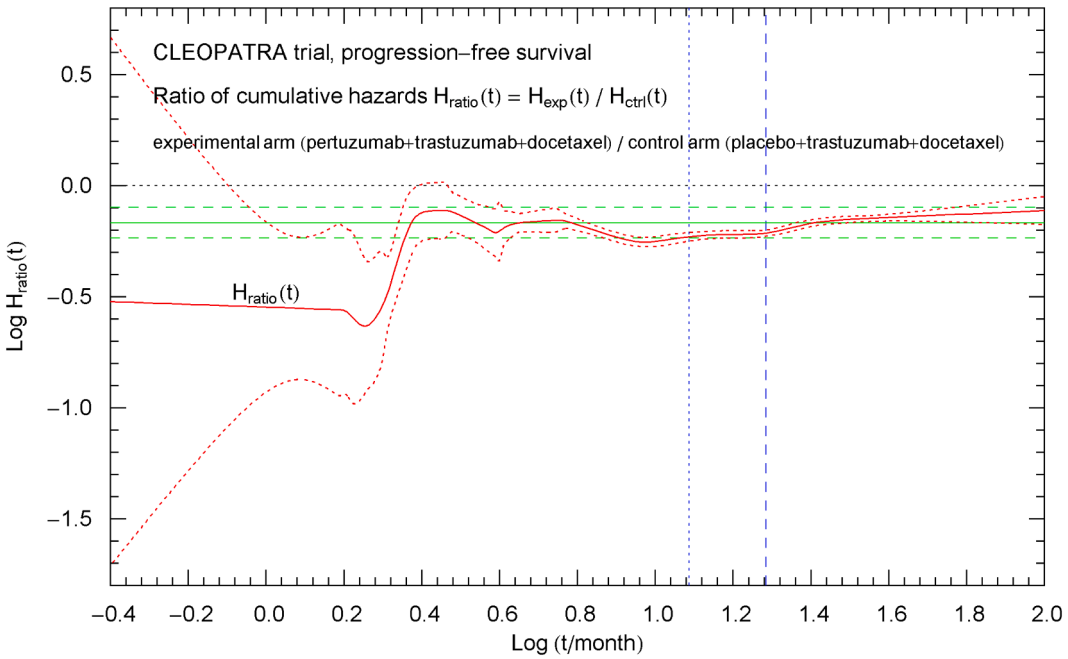

Fig. 15. Time evolution of the cumulative-hazard ratio of the CLEOPATRA trial [12], for PFS. Depicted is the ratio $H_{\rm exp}(t)/H_{\rm ctrl}(t)$ (red solid curve, cf. Section 4) of the cumulative hazard functions of the experimental arm $H_{\rm exp}(t)$ and control arm $H_{\rm ctrl}(t)$, cf. Figs. 12,14. The red dotted curves are the boundaries of the 95% confidence band of the ratio. For comparison, the green solid horizontal line shows the constant hazard ratio of the Cox regression model, and the green dashed horizontal lines depict the 95% confidence interval thereof, cf. Ref. [12]. The blue dotted and dashed vertical lines indicate the median PFS time of the control and experimental arm, respectively, cf. Table 4.

Table 1
Fitting parameters of the survival functions of the IBCSG 22–00 and CLEOPATRA trials, cf. Section 2. The survival functions are modeled as exponentiated broken power-law distributions $S(t) = \exp(-H(t))$, with the cumulative hazard function $H(t) = a_0 t^{\alpha_0}\prod_{k=1}^{n}\left(1+(t/10^{c_k})^{\beta_k/|\eta_k|}\right)^{\eta_k}$. The parameters a_0, α_0, $(c_k,\beta_k,\eta_k)_{k=1,\ldots,n}$, regressed from KM data points (extracted from uncensored and censored event times [1]), are recorded in this table. Listed are the parameters of $S(t)$ and $H(t)$ for disease-free survival (DFS, IBCSG 22–00) and overall and progression-free survival (OS and PFS, CLEOPATRA), in each case for the experimental (exp.) and control (ctrl.) arm. The regression is based on the least-squares functional χ^2_S in Eq. (2.2). The value of this functional at the regressed parameters and the degrees of freedom (dof: number N of data points in the least-squares functional minus number of fitting parameters) are recorded as well. Also listed are the standard error (SE) and the coefficient of determination R^2. Figs. 1–4 (for DFS, IBCSG 22–00) and Figs. 6–9 (for OS, CLEOPATRA) and Figs. 11–14 (for PFS, CLEOPATRA) show Log–Lin and Log–Log plots of the regressed survival and cumulative hazard functions, including residual panels.

	IBCSG DFS, exp.	IBCSG DFS, ctrl.	CLEO OS, exp.	CLEO OS, ctrl.	CLEO PFS, exp.	CLEO PFS, ctrl.
a_0	0.0292413	0.359769	0.0101044	0.0153306	0.00473555	0.016675
α_0	1.00031	3.04814	0.688387	1.30891	1.02636	1.08889
c_1	– 0.232585	– 0.433434	1.04462	0.0653435	0.259462	0.205931
β_1	2.24455	1.31179	1.68293	0.887251	6.00997	2.38821
η_1	0.173507	– 0.0309817	0.0937996	– 0.0272857	0.200194	0.0299254
c_2	– 0.123937	– 0.157079	1.22529	0.79127	0.371478	0.310896
β_2	1.80557	0.572184	1.01754	1.39472	6.505	2.96352
η_2	– 0.0410148	– 0.0118065	– 0.0379746	0.0371325	– 0.199805	– 0.0501654
c_3	0.286986	0.108293	1.43121	0.972601	0.591855	0.471866
β_3	0.803633	1.77186	0.10715	0.27074	3.43512	0.848373
η_3	– 0.0348929	0.0113317	–0.00431582	–0.00905803	0.0301537	0.0108402
c_4	0.559319	0.166884	–	1.40631	0.60713	0.762259
β_4	0.307975	2.44264	–	0.350323	2.49009	0.699571
η_4	– 0.0121684	– 0.322015	–	– 0.0131034	– 0.0793856	0.0122501
c_5	–	–	–	–	1.11738	0.960159
β_5	–	–	–	–	0.402613	0.938036
η_5	–	–	–	–	– 0.0379167	– 0.0538269
c_6	–	–	–	–	1.40552	1.28797
β_6	–	–	–	–	0.451349	0.582016
η_6	–	–	–	–	– 0.0267739	– 0.0314647
R^2	0.999410	0.999252	0.998721	0.998827	0.998681	0.997659
SE	0.00165	0.00227	0.00476472	0.0058767	0.00864852	0.0125486
χ^2_S	1.833	2.635	7.63653	12.6511	34.484	68.767
dof	104 – 14	137 – 14	162 – 11	195 – 14	192 – 20	211 – 20

Table 2
Selected percentiles of the regressed survival functions of the experimental (exp.) and control (ctrl.) arm of the IBCSG 22–00 trial, for disease-free survival (DFS), cf. Ref. [11]. The listed $t_{k/100}$ denotes the time at which disease recurrence in k% of patients is detected, the zero point of the time axis being the time of randomization. The k-th percentile $t_{k/100}$ is obtained by solving $1-k/100 = S(t_{k/100})$, where $S(t)$ is the regressed analytic survival function, cf. Figs. 1,3. The 95% confidence intervals (CIs) of the percentiles are also indicated.

k %	IBCSG DFS, exp. $t_{k/100}$ [yrs]	IBCSG DFS, exp. 95% CI	IBCSG DFS, ctrl. $t_{k/100}$ [yrs]	IBCSG DFS, ctrl. 95% CI
5	0.917	0.849 - 0.984	0.696	0.640 - 0.752
10	1.51	1.46 - 1.56	1.34	1.28 - 1.40
15	2.19	1.99 - 2.40	1.79	1.70 - 1.88
20	3.66	3.19 - 4.12	3.01	2.80 - 3.21
25	7.77	5.56 - 9.98	5.02	4.72 - 5.31
30	–	–	7.76	6.96 - 8.55

Table 3
Disease-free survival rate $S(t)$ and cumulative-hazard ratio $H_{\text{ratio}} = H_{\text{exp}}(t)/H_{\text{ctrl}}(t)$ (for DFS) of the experimental (exp.) and control (ctrl.) arm of the IBCSG 22–00 trial [11] at selected times (starting from randomization), cf. Sections 3 and 4. Also indicated in the adjacent columns are the 95% confidence intervals (CIs) as well as the p-value of $H_{\text{ratio}}(t)$. The time-varying hazard ratio and its 95% confidence band are depicted in Fig. 5.

t [yrs]	IBCSG DFS, exp. $S(t)$[%]	IBCSG DFS, exp. 95% CI	IBCSG DFS, ctrl. $S(t)$[%]	IBCSG DFS, ctrl. 95% CI	IBCSG DFS $H_{\text{ratio}}(t)$	IBCSG DFS 95% CI	IBCSG DFS p-value
0.5	98.5	97.8 - 99.2	97.1	96.7 - 97.5	0.512	0.311 - 0.844	0.00862
1	94.4	93.9 - 94.8	92.5	92.0 - 93.0	0.748	0.669 - 0.838	< 0.0001
2	85.9	85.2 - 86.7	83.7	83.2 - 84.3	0.854	0.797 - 0.915	< 0.0001
3	82.0	81.1 - 82.8	80.0	79.4 - 80.6	0.891	0.838 - 0.948	0.000235
4	79.4	78.0 - 80.7	77.3	76.8 - 77.9	0.899	0.830 - 0.974	0.00890
5	78.0	77.0 - 79.0	75.0	74.4 - 75.7	0.867	0.817 - 0.920	< 0.0001
6	76.8	75.7 - 77.8	73.0	72.2 - 73.9	0.841	0.789 - 0.897	< 0.0001
7	75.7	74.2 - 77.3	71.2	70.2 - 72.3	0.820	0.752 - 0.894	< 0.0001
8	74.8	72.6 - 76.9	69.6	68.3 - 70.9	0.802	0.717 - 0.897	0.000108
9	73.9	71.2 - 76.7	68.1	66.6 - 69.7	0.786	0.686 - 0.901	0.000544
10	73.2	69.9 - 76.5	66.7	64.9 - 68.5	0.773	0.659 - 0.906	0.00148

Table 4
Selected percentiles of the regressed survival functions of the experimental (exp.) and control (ctrl.) arm of the CLEOPATRA trial, for overall (OS) and progression-free survival (PFS). The listed $t_{k/100}$ denotes the time at which death or disease progression in k% of patients is observed, cf. Section 3. The 95% confidence intervals (CIs) of the percentiles are also indicated.

k %	CLEO OS, exp. $t_{k/100}$ [months]	CLEO OS, exp. 95% CI	CLEO OS, ctrl. $t_{k/100}$ [months]	CLEO OS, ctrl. 95% CI	CLEO PFS, exp. $t_{k/100}$ [months]	CLEO PFS, exp. 95% CI	CLEO PFS, ctrl. $t_{k/100}$ [months]	CLEO PFS, ctrl. 95% CI
5	10.3	8.89 - 11.6	7.32	6.38 - 8.25	2.43	2.17 - 2.68	1.91	1.67 - 2.16
10	14.8	14.4 - 15.3	11.2	10.7 - 11.6	4.84	4.59 - 5.09	3.68	3.47 - 3.90
15	18.6	17.7 - 19.5	14.8	14.4 - 15.1	6.50	6.28 - 6.71	5.07	4.75 - 5.38
20	23.5	22.7 - 24.2	18.1	17.8 - 18.5	8.05	7.86 - 8.24	6.18	6.06 - 6.30
25	28.5	27.6 - 29.3	21.4	20.9 - 21.8	9.57	9.33 - 9.82	6.99	6.89 - 7.09
30	33.8	32.9 - 34.6	24.6	24.1 - 25.2	11.1	10.8 - 11.4	7.77	7.64 - 7.90
40	45.1	44.1 - 46.0	32.8	32.3 - 33.4	14.6	14.2 - 15.1	9.51	9.30 - 9.71
50	57.6	55.3 - 59.8	42.4	41.7 - 43.0	19.3	18.9 - 19.7	12.2	12.0 - 12.5
60	72.0	67.7 - 76.3	53.5	52.2 - 54.8	25.4	24.7 - 26.2	15.7	15.4 - 16.0
70	–	–	–	–	38.3	36.9 - 39.6	20.9	20.1 - 21.8
75	–	–	–	–	48.0	45.2 - 50.8	26.8	25.7 - 27.9
80	–	–	–	–	61.0	55.2 - 66.9	35.3	33.9 - 36.7

Table 5
Overall survival rate $S(t)$ and cumulative-hazard ratio $H_{\rm ratio}=H_{\rm exp}(t)/H_{\rm ctrl}(t)$ for OS of the CLEOPATRA trial [12] at selected times, calculated from the regressed survival and hazard functions of the experimental (exp.) and control (ctrl.) arm, cf. Figs. 6–9 and Sections 3 and 4. Also indicated are the 95% confidence intervals (CIs) and the p-value of $H_{\rm ratio}(t)$. The hazard ratio for OS and its confidence band are depicted in Fig. 10.

t [months]	CLEO OS, exp. $S(t)$[%]	CLEO OS, exp. 95% CI	CLEO OS, ctrl. $S(t)$[%]	CLEO OS, ctrl. 95% CI	CLEO OS $H_{\rm ratio}(t)$	CLEO OS 95% CI	CLEO OS p-value
6	96.6	96.1 - 97.0	96.3	95.3 - 97.3	0.921	0.685 - 1.24	0.583
12	93.7	93.0 - 94.4	88.9	88.3 - 89.5	0.553	0.484 - 0.633	<0.0001
18	85.6	84.8 - 86.5	80.2	79.7 - 80.7	0.703	0.655 - 0.754	<0.0001
24	79.5	78.6 - 80.3	71.0	70.1 - 71.8	0.670	0.633 - 0.709	<0.0001
30	73.5	72.7 - 74.4	63.2	62.5 - 63.9	0.671	0.640 - 0.702	<0.0001
36	68.0	67.3 - 68.7	56.5	56.0 - 57.1	0.677	0.656 - 0.699	<0.0001
42	62.6	62.0 - 63.3	50.4	49.8 - 51.0	0.682	0.663 - 0.702	<0.0001
48	57.5	56.6 - 58.5	44.7	43.9 - 45.6	0.687	0.661 - 0.714	<0.0001
54	52.7	51.3 - 54.1	39.6	38.5 - 40.7	0.691	0.656 - 0.728	<0.0001
60	48.2	46.3 - 50.1	35.0	33.6 - 36.3	0.695	0.651 - 0.742	<0.0001
66	44.0	41.6 - 46.3	30.8	29.2 - 32.4	0.698	0.646 - 0.755	<0.0001
72	40.0	37.3 - 42.7	27.1	25.3 - 28.9	0.701	0.641 - 0.768	<0.0001

Table 6
Progression-free survival rate $S(t)$ and cumulative-hazard ratio $H_{\rm ratio}=H_{\rm exp}(t)/H_{\rm ctrl}(t)$ for PFS of the CLEOPATRA trial [12] at selected times, calculated from the regressed survival and hazard functions of the experimental (exp.) and control (ctrl.) arm, cf. Figs. 11–14. Also indicated are the 95% confidence intervals (CIs) and the p-value of $H_{\rm ratio}(t)$. The time-varying hazard ratio for PFS and its confidence band are depicted in Fig. 15.

t [months]	CLEO PFS, exp. $S(t)$[%]	CLEO PFS, exp. 95% CI	CLEO PFS, ctrl. $S(t)$[%]	CLEO PFS, ctrl. 95% CI	CLEO PFS $H_{\rm ratio}(t)$	CLEO PFS 95% CI	CLEO PFS p-value
6	86.5	85.8 - 87.3	81.0	80.3 - 81.8	0.687	0.641 - 0.738	<0.0001
12	67.2	66.2 - 68.2	50.8	49.9 - 51.7	0.586	0.560 - 0.614	<0.0001
18	52.6	51.8 - 53.4	34.6	33.6 - 35.6	0.606	0.585 - 0.628	<0.0001
24	41.9	41.0 - 42.9	27.1	26.2 - 28.1	0.667	0.642 - 0.692	<0.0001
30	35.6	34.5 - 36.6	22.9	22.3 - 23.6	0.701	0.677 - 0.726	<0.0001
36	31.4	30.6 - 32.2	19.7	19.0 - 20.4	0.713	0.691 - 0.735	<0.0001
42	27.9	27.0 - 28.8	17.1	16.2 - 18.0	0.722	0.694 - 0.751	<0.0001
48	25.0	23.8 - 26.2	15.0	13.9 - 16.0	0.730	0.693 - 0.769	<0.0001
54	22.5	20.9 - 24.1	13.2	12.0 - 14.4	0.737	0.690 - 0.787	<0.0001
60	20.3	18.5 - 22.2	11.7	10.4 - 13.0	0.743	0.687 - 0.804	<0.0001

Data availability

The data files containing event times and censoring information can be downloaded from cancertrials.io.

CRediT authorship contribution statement

Roman Tomaschitz: Conceptualization, Investigation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

My thanks to two anonymous referees for their comments and suggestions, which have greatly helped to improve an earlier draft of this paper.

Appendix A. Polygonal chain limits of multiply broken power laws and exponentials

We will demonstrate that a multiply broken power-law density has a unique polygonal chain limit in Log−Log coordinates, cf. Section 2. Conversely, every polygonal chain defining a single-valued function on the real axis is a limit distribution of a multiply broken power-law density. The proof indicated below is by explicit construction of the polygonal chain, starting with a broken power-law density, and vice versa. That is, we establish an explicit invertible relation between the parameters of the broken power law and the polygonal chain, cf. Sections A.1 and A.2. In Section A.3, the polygonal chain limit of multiply broken exponentials in Log−Lin coordinates is studied, in like manner.

A.1. Parametrization of polygonal chains

We consider a sequence of points $P_i=(\widehat{b}_i,y_i)$, $i=0,\ldots,n+1$, with abscissas $\widehat{b}_i<\widehat{b}_{i+1}$ in ascending order, defining a polygonal chain in the Cartesian (x,y) plane. The polygonal edges are the line segments

$$L_i:\ y=k_ix+d_i,\ \widehat{b}_i\le x\le\widehat{b}_{i+1},\ i=0,\ldots,n, \tag{A.1}$$

with slope and vertical intercept

$$k_i = \frac{y_{i+1}-y_i}{\widehat{b}_{i+1}-\widehat{b}_i},\ d_i = \frac{\widehat{b}_{i+1}y_i - \widehat{b}_i y_{i+1}}{\widehat{b}_{i+1}-\widehat{b}_i}. \tag{A.2}$$

The polygonal vertices are $P_i = (\widehat{b}_i, y_i)$, $i = 1,\ldots,n$, excluding the endpoints P_0 and P_{n+1} of the chain. We may write the vertices as $P_i = (\widehat{b}_i, k_i\widehat{b}_i + d_i) = (\widehat{b}_i, k_{i-1}\widehat{b}_i + d_{i-1})$, $i = 1,\ldots,n$, since adjacent edges L_i join at the vertices, and the endpoints read $P_0 = (\widehat{b}_0, k_0\widehat{b}_0 + d_0)$, $P_{n+1} = (\widehat{b}_{n+1}, k_n\widehat{b}_{n+1} + d_n)$. The vertical intercepts d_i of the straight lines defining the edges thus satisfy $d_i = (k_{i-1} - k_i)\widehat{b}_i + d_{i-1}$, $i = 1,\ldots,n$. Iterating this, we find

$$d_i = d_0 + \sum_{j=1}^{i}(k_{j-1} - k_j)\widehat{b}_j,\ i = 0,\ldots,n. \tag{A.3}$$

The polygonal chain is thus defined by specifying the vertical intercept d_0 of the first edge L_0, the slope parameters k_i of the edges, $i = 0,\ldots,n$, and the abscissas $\widehat{b}_i$, $i = 1,\ldots,n$, of the vertices. d_0 and k_i are real parameters, the slopes of adjacent edges do not coincide, $k_i \neq k_{i+1}$, and the abscissas $\widehat{b}_i$ of the vertices satisfy $\widehat{b}_i < \widehat{b}_{i+1}$.

Using Iverson brackets (which evaluate to one if the inequality/equality in them is satisfied and to zero otherwise), the polygonal chain can be written as

$$y = \sum_{i=0}^{n} y_i(x),\ \ y_i(x) = (k_i x + d_i)[\widehat{b}_i \le x < \widehat{b}_{i+1}], \tag{A.4}$$

with intercepts d_i in (A.3). The first and last polygonal edge L_0 and L_n can be extended to infinity by putting $\widehat{b}_0 = -\infty$ and $\widehat{b}_{n+1} = +\infty$ in the bracket (without changing the k_i and d_i in (A.2)), in which case the extended polygonal chain defines a one-valued function on the real axis.

A.2. Polygonal limits of multiply broken power laws in Log–Log coordinates

We consider the multiply broken power-law distribution

$$H(t) = a_0 t^{\alpha_0}\prod_{k=1}^{n}\left(1 + (t/b_k)^{\beta_k/\eta_k}\right)^{s_k\eta_k}, \tag{A.5}$$

defined on the positive real axis. The parameters are the positive exponents β_k and η_k, real exponent α_0, and positive amplitudes a_0 and b_k. The exponents s_k can take the values ± 1. The number n of factors in the product is adaptable to the data set to which the density is fitted, and the factors in (A.5) are ordered by increasing amplitude $b_k < b_{k+1}$.

Density $H(t)$ in (A.5) is composed of factors $p_k(t) = \left(1 + (t/b_k)^{\beta_k/\eta_k}\right)^{s_k\eta_k}$, $k = 1,\ldots,n$, which can be split as

$$p_k(t) = [t<b_k]\left(1 + (t/b_k)^{\beta_k/\eta_k}\right)^{s_k\eta_k} + [t=b_k]2^{s_k\eta_k} + [b_k<t](t/b_k)^{\beta_k s_k}\left(1 + (t/b_k)^{-\beta_k/\eta_k}\right)^{s_k\eta_k}, \tag{A.6}$$

where Iverson brackets are used as in (A.4).

The $\eta_k \to 0$ limit of (A.6) reads

$$p_k(t) \sim [t<b_k] + [b_k \le t](t/b_k)^{\beta_k s_k}. \tag{A.7}$$

The factors $p_k(t)$ in the product $H(t) = a_0 t^{\alpha_0}\prod_{k=1}^{n} p_k(t)$ are ordered, $b_k < b_{k+1}$, and we also define $b_0 = 0$ and $b_{n+1} = +\infty$. By substituting the asymptotic factors (A.7) and multiplying out the product, we find

$$H(t) \sim \sum_{k=0}^{n}[b_k \le t < b_{k+1}]a_0 t^{\alpha_0}\prod_{i=1}^{k}(t/b_i)^{\beta_i s_i}. \tag{A.8}$$

This limit can be written as a finite series of power-law segments,

$$H(t) \sim \sum_{i=0}^{n} q_i(t),\ q_i(t) = [b_i \le t < b_{i+1}]a_0\left(\prod_{j=1}^{i} b_j^{-s_j\beta_j}\right)t^{\alpha_0 + \sum_{l=1}^{i} s_l\beta_l}. \tag{A.9}$$

The power-law segments $q_i(t)$ are mapped onto the polygonal edges $y_i(x) = (k_i x + d_i)[\widehat{b}_i \le x < \widehat{b}_{i+1}]$ in (A.4) by introducing (decadic) Log–Log coordinates, identifying $t = 10^x$, $y = \mathrm{Log}\, H$, so that $y_i = \mathrm{Log}\, q_i$ with interval boundaries in (A.4) and (A.9) related by $\widehat{b}_i = \mathrm{Log}\, b_i$. Explicitly,

$$\mathrm{Log}\, q_i(t) = [\mathrm{Log}\, b_i \le x < \mathrm{Log}\, b_{i+1}]\left(\mathrm{Log}\left(a_0\prod_{j=1}^{i} b_j^{-s_j\beta_j}\right) + x\left(\alpha_0 + \sum_{j=1}^{i} s_j\beta_j\right)\right). \tag{A.10}$$

Thus the parameters in (A.4) and (A.10) can be identified as

$$k_i = \alpha_0 + \sum_{j=1}^{i} s_j\beta_j, \tag{A.11}$$

$$d_i = \mathrm{Log}\, a_0 - \sum_{j=1}^{i} s_j\beta_j \mathrm{Log}\, b_j, \tag{A.12}$$

with $i = 0,\ldots,n$. The abscissas of the polygonal vertices in Cartesian (x,y) coordinates are $\widehat{b}_i = \mathrm{Log}\, b_i$, cf. Section A.1. This defines the parameters of the polygonal chain Eq. (A.4) once the multiply broken power law $H(t)$ in (A.5) is specified. (The positive η_j exponents of $H(t)$ do not enter in (A.11) and (A.12), since the chain is defined by the $\eta_j \to 0$ limit of $H(t)$, cf. (A.5).)

We may invert the linear Eqs. (A.11), (A.12) for the variables α_0, $s_j\beta_j$ and $\mathrm{Log}\, b_j$. Inverting (A.11) by way of $k_i = k_{i-1} + s_i\beta_i$ yields

$$\alpha_0 = k_0,\ s_i\beta_i = k_i - k_{i-1}, \tag{A.13}$$

with $i = 1,\ldots,n$. This gives the real exponent α_0, the positive exponents $\beta_i = |k_i - k_{i-1}|$ and the $\pm$ signs $s_i = \mathrm{sign}(k_i - k_{i-1})$ of $H(t)$ in (A.5).

Eq. (A.12) can be solved for $\mathrm{Log}\, a_0$ and $\mathrm{Log}\, b_i$. To this end, we substitute $s_j\beta_j = k_j - k_{j-1}$ and use $d_i = d_{i-1} - (k_i - k_{i-1})\mathrm{Log}\, b_i$ (compare also to (A.3)) to find

$$\mathrm{Log}\, a_0 = d_0,\ \mathrm{Log}\, b_i = -\frac{d_i - d_{i-1}}{k_i - k_{i-1}}, \tag{A.14}$$

with $i = 1,\ldots,n$. The denominators in (A.14) cannot vanish, since adjacent edges of the polygonal chain have different slopes. This defines the positive amplitudes a_0 and b_i of density $H(t)$ in (A.5).

Thus the parameters a_0, α_0, b_i, β_i, s_i, $i = 1,\ldots,n$, of the multiply broken power-law density $H(t)$ in (A.5) can unambiguously be determined from a prescribed polygonal chain in Log–Log coordinates $x = \mathrm{Log}\, t$, $y = \mathrm{Log}\, H$, cf. Section A.1. As mentioned, the positive exponents η_i of $H(t)$ are not determined by the chain, since the chain is the $\eta_i \to 0$ limit of $H(t)$, cf. (A.7).

By using small but finite η_i exponents in the broken power-law density (A.5), the polygonal chain (emerging in the $\eta_i \to 0$ limit in Log–Log coordinates) is approximated by a smooth curve. This smoothed polygon (with rounded vertices and curved edges) represents an analytic function. If the η_i exponents become moderate (in the least-squares iteration), the regressed curve bears little resemblance to its polygonal $\eta_i \to 0$ limit.

To find initial values for the multiparametric least-squares functional Eq. (2.2), one can start with a polygonal approximation of the empirical curve defined by the data set for the cumulative hazard $H(t)$. This can be done by visually fitting the polygonal edges successively, cf. Ref. [35]. An automated way to obtain a polygonal approximation of $H(t)$ with a limited number of vertices is to use the Ramer–Douglas–Peucker algorithm, cf. Refs. [36,37], which constructs a polygonal chain that is epsilon-close to the data points, the number of vertices depending on the choice of epsilon (allowed maximum distance). By slightly modifying this divide-and-rule algorithm, one can prescribe the number of vertices of the approximating polyline instead of the maximum distance to the data points.

A.3. Polygonal limits of multiply broken exponential densities in Log–Lin coordinates

Multiply broken exponentials are defined analogously to the broken power-law densities (A.5),

$$H(t) = a_0 \mathrm{e}^{\alpha_0 t} \prod_{k=1}^{n} \left(1 + \mathrm{e}^{(\beta_k/\eta_k)(t-b_k)}\right)^{s_k\eta_k}, \tag{A.15}$$

with positive exponents β_k, η_k, real exponent α_0, positive amplitude a_0 and real amplitudes b_k. The exponents s_k can take the values ± 1. The factors are ordered by increasing amplitude, $b_k < b_{k+1}$, and we also define $b_0 = -\infty$ and $b_{n+1} = +\infty$. These densities are best depicted in decadic Log–Lin coordinates, where simple exponentials appear as straight lines. Evidently, $H(t)$ in (A.15) is obtained from $H(t)$ in (A.5) by the substitutions $t \to \mathrm{e}^t$, $b_k \to \mathrm{e}^{b_k}$. A similar density is used as cumulative hazard function in life-table modeling, cf. Ref. [19]. The subsequent reasoning is analogous to Section A.2, with minor changes.

Density $H(t)$ in (A.15) is composed of factors $p_k(t) = \left(1 + \mathrm{e}^{(\beta_k/\eta_k)(t-b_k)}\right)^{s_k\eta_k}$, which can be split as

$$p_k(t) = [t < b_k]\left(1 + \mathrm{e}^{(\beta_k/\eta_k)(t-b_k)}\right)^{s_k\eta_k} + [t = b_k]2^{s_k\eta_k} + [b_k < t]\mathrm{e}^{\beta_k s_k(t-b_k)}\left(1 + \mathrm{e}^{-(\beta_k/\eta_k)(t-b_k)}\right)^{s_k\eta_k}, \tag{A.16}$$

where we use Iverson brackets as in (A.4). Performing the $\eta_k \to 0$ limit in (A.16) yields

$$p_k(t) \sim [t < b_k] + [b_k \le t]\mathrm{e}^{\beta_k s_k(t-b_k)}. \tag{A.17}$$

The factors $p_k(t)$ in the product $H(t) = a_0\mathrm{e}^{\alpha_0 t}\prod_{k=1}^{n} p_k(t)$ are ordered by increasing amplitude, $b_k < b_{k+1}$. The $\eta_k \to 0$ limit of the broken exponential density $H(t)$ in (A.15) is readily found by substituting $t \to \mathrm{e}^t$, $b_i \to \mathrm{e}^{b_i}$ into (A.8) (leaving the Iverson brackets unchanged),

$$H(t) \sim \sum_{k=0}^{n} [b_k \le t < b_{k+1}] a_0 \mathrm{e}^{\alpha_0 t} \prod_{i=1}^{k} \mathrm{e}^{\beta_i s_i (t-b_i)}. \tag{A.18}$$

This can be written as a finite series of exponential segments,

$$H(t) \sim \sum_{i=0}^{n} q_i(t),\ q_i(t) = [b_i \le t < b_{i+1}] a_0 \mathrm{e}^{t\left(\alpha_0 + \sum_{j=1}^{i}\beta_j s_j\right) - \sum_{j=1}^{i} b_j\beta_j s_j}. \tag{A.19}$$

The exponentials $q_i(t)$ are mapped onto the polygonal edges $y_i(x) = (k_i x + d_i)[\widehat{b}_i \le x < \widehat{b}_{i+1}]$ in (A.4) by introducing decadic Log–Lin coordinates, identifying $x = t$, $y = \mathrm{Log}\, H$, so that $y_i = \mathrm{Log}\, q_i$ and the interval boundaries in (A.4) and (A.19) coincide, $\widehat{b}_i = b_i$. In particular,

$$\mathrm{Log}\, q_i(t) = [b_i \le x < b_{i+1}]\left(\mathrm{Log}\, a_0 - \sum_{j=1}^{i} b_j\beta_j s_j \mathrm{Log}\, \mathrm{e} + x\left(\alpha_0 + \sum_{j=1}^{i}\beta_j s_j\right)\mathrm{Log}\, \mathrm{e}\right), \tag{A.20}$$

with Log e $= 1/\log 10$. The parameters in (A.4) and (A.20) can be identified as

$$k_i = \alpha_0 \text{Log e} + \sum_{j=1}^{i} s_j \beta_j \text{Log e}, \tag{A.21}$$

$$d_i = \text{Log } a_0 - \sum_{j=1}^{i} s_j \beta_j b_j \text{Log e}, \tag{A.22}$$

with $i = 0, \ldots, n$. The abscissas of the polygonal vertices in Cartesian coordinates read $\widehat{b}_i = b_i$, cf. Section A.1. This defines the parameters of the polygonal chain Eq. (A.4) once the multiply broken exponential $H(t)$ in (A.15) is specified. The η_j exponents do not enter in (A.21) and (A.22), since the polygonal chain is defined by the $\eta_j \to 0$ limit of $H(t)$, cf. (A.17).

The linear equations (A.21) and (A.22) can be inverted for the variables α_0, $s_j\beta_j$ and b_j. Inverting (A.21) by way of $k_i = k_{i-1} + s_i\beta_i \text{Log e}$ yields

$$\alpha_0 = k_0/\text{Log e}, s_i\beta_i = (k_i - k_{i-1})/\text{Log e}, \tag{A.23}$$

with $i = 1, \ldots, n$. This determines the real exponent α_0, the positive exponents β_i and the $\pm$ signs s_i of $H(t)$ in (A.15).

Eq. (A.22) can be solved for Log a_0 and b_i. To this end, we substitute $s_j\beta_j = (k_j - k_{j-1})/\text{Log e}$ and use $d_i = d_{i-1} - (k_i - k_{i-1})b_i$ to find

$$\text{Log } a_0 = d_0, \; b_i = -\frac{d_i - d_{i-1}}{k_i - k_{i-1}}, \tag{A.24}$$

with $i = 1, \ldots, n$. The denominators in (A.24) cannot vanish, since adjacent polygonal edges have different slopes. This gives the positive amplitude a_0 and the real amplitudes b_i of $H(t)$ in (A.15).

Thus the parameters a_0, α_0, b_i, β_i, s_i, $i = 1, \ldots, n$, of the multiply broken exponential density $H(t)$ in (A.15) can unambiguously be determined from a prescribed polygonal chain (approximating the data set) in Log−Lin coordinates $x = t, y = \text{Log } H$, cf. Section A.1. The positive exponents η_i of $H(t)$ are not determined by the chain, since the chain is the $\eta_i \to 0$ limit of $H(t)$, cf. (A.17).

The choice between broken power-laws and broken exponential densities can readily be made by plotting the data in Log−Log and Log−Lin coordinates and inspecting the slope of the empirical curve defined by the data points. For instance, if the KM data of the cumulative hazard in Figs. 2,4,7,9,12,14 are plotted in Log−Lin coordinates, they exhibit a steep initial slope, which would require large exponents if the hazard function is modeled as broken exponential (A.15). In this case, one uses Log−Log coordinates and the multiply broken power law (A.5) as cumulative hazard. In contrast, employing a multiply broken power law for the intermediate age range in life tables (Gompertz regime [24]) would require large power-law exponents in (A.5) for the steep Log−Log slope. In this case, one uses exponential modeling of the hazard, which is done in Log−Lin coordinates [19].

References

[1] http://cancertrials.io.

[2] D. Plana, G. Fell, B.M. Alexander, A.C. Palmer, P.K. Sorger, Cancer patient survival can be parametrized to improve trial precision and reveal time-dependent therapeutic effects, Nat. Commun. 13 (2022) 873, https://doi.org/10.1038/s41467-022-28410-9.

[3] D. Plana, G. Fell, B.M. Alexander, A.C. Palmer, P.K. Sorger, Imputed individual participant data from oncology clinical trials, Synapse (2021), https://doi.org/10.7303/SYN25813713.

[4] J.D. Kalbfleisch, R.L. Prentice, The Statistical Analysis of Failure Time Data, 2nd ed., Wiley, Hoboken, NJ, 2002.

[5] E.T. Lee, J.W. Wang, Statistical Methods for Survival Data Analysis, 3rd ed, Wiley, Hoboken, NJ, 2003.

[6] D. Collett, Modelling Survival Data in Medical Research, 3rd ed, CRC Press, Boca Raton, FL, 2015.

[7] J.F. Lawless, Statistical Models and Methods for Lifetime Data, 2nd ed, Wiley, Hoboken, NJ, 2003.

[8] J. Ma, P. Dhiman, C. Qi, G. Bullock, M. van Smeden, R.D. Riley, G.S. Collins, Poor handling of continuous predictors in clinical prediction models using logistic regression: a systematic review, J. Clin. Epidemiol. 161 (2023) 140–151, https://doi.org/10.1016/j.jclinepi.2023.07.017.

[9] S.J. Tingle, G. Kourounis, S. Elliot, E.M. Harrison, Non-linear regression modelling for medical professionals; making curved paths straightforward, Postgrad. Med. J. (2025) qgaf183, https://doi.org/10.1093/postmj/qgaf183.

[10] P.C. Austin, J. Fang, D.S. Lee, Using fractional polynomials and restricted cubic splines to model non-proportional hazards or time-varying covariate effects in the Cox regression model, Stat. Med. 41 (2022) 612–624, https://doi.org/10.1002/sim.9259.

[11] M. Colleoni, K.P. Gray, S. Gelber, I. Láng, B. Thürlimann, et al., Low-dose oral cyclophosphamide and methotrexate maintenance for hormone receptor–negative early breast cancer: International Breast Cancer Study Group Trial 22-00, J. Clin. Oncol. 28 (2016) 3400–3410, https://doi.org/10.1200/JCO.2015.65.6595.

[12] S.M. Swain, J. Baselga, S.-B. Kim, J. Ro, V. Semiglazov, M. Campone, E. Ciruelos, J.-M. Ferrero, A. Schneeweiss, S. Heeson, E. Clark, G. Ross, M.C. Benyunes, J. Cortés, Pertuzumab, trastuzumab, and docetaxel in HER2-positive metastatic breast cancer, N. Engl. J. Med. 372 (2015) 724–734, https://doi.org/10.1056/NEJMoa1413513.

[13] J. Baselga, J. Cortés, S.-B. Kim, S.-A. Im, R. Hegg, Y.-H. Im, L. Roman, J.L. Pedrini, T. Pienkowski, A. Knott, E. Clark, M.C. Benyunes, G. Ross, S.M. Swain, Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer, N. Engl. J. Med. 366 (2012) 109–119, https://doi.org/10.1056/NEJMoa1113216.

[14] S.M. Swain, S.-B. Kim, J. Cortés, J. Ro, V. Semiglazov, M. Campone, E. Ciruelos, J.-M. Ferrero, A. Schneeweiss, A. Knott, E. Clark, G. Ross, M.C. Benyunes, J. Baselga, Pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA study): overall survival results from a randomised, double-blind, placebo-controlled, phase 3 study, Lancet Oncol. 14 (2013) 461–471, https://doi.org/10.1016/S1470-2045(13)70130-X.

[15] P. Guyot, A. Ades, M.J. Ouwens, N.J. Welton, Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan−Meier survival curves, BMC Med. Res. Methodol. 12 (2012) 9, https://doi.org/10.1186/1471-2288-12-9.

[16] G. Fell, R.A. Redd, A.M. Vanderbeek, R. Rahman, B. Louv, et al., KMDATA: a curated database of reconstructed individual patient-level data from 153 oncology clinical trials, Database baab037 (2021), https://doi.org/10.1093/database/baab037.

[17] A.C. Palmer, B. Izar, H. Hwangbo, P.K. Sorger, Predictable clinical benefits without evidence of synergy in trials of combination therapies with immune-checkpoint inhibitors, Clin. Cancer Res. 28 (2022) 368–377, https://doi.org/10.1158/1078-0432.CCR-21-2275.

[18] R. Tomaschitz, Multiply broken power-law densities as survival functions: an alternative to Pareto and lognormal fits, Physica A 541 (2020) 123188, https://doi.org/10.1016/j.physa.2019.123188.

[19] R. Tomaschitz, Nonlinear multiparametric modeling of life-table data with adaptive distributions: time evolution of hazard ratios, Eur. Phys. J. Plus 139 (2024) 982, https://doi.org/10.1140/epjp/s13360-024-05746-3.

[20] R. Tomaschitz, Modeling heat-capacity data with multiparametric hazard functions, Appl. Phys. A 131 (2025) 353, https://doi.org/10.1007/s00339-025-08423-z.

[21] Wolfram Research, Inc., Mathematica,® version 14.3, Champaign, IL, 2025, https://www.wolfram.com/mathematica/.

[22] T.O. Kvålseth, Cautionary note about R^2, Am. Stat. 39 (1985) 279–285, https://doi.org/10.1080/00031305.1985.10479448.

[23] P.H. Richter, Estimating errors in least-squares fitting, in: TDA Progress Report, 42-122, 1995. https://ipnpr.jpl.nasa.gov/progress_report/42-122/122E.pdf.

[24] A.W. Marshall, I. Olkin, Life Distributions, Springer, New York, 2007.

[25] M.A. Hernán, The hazards of hazard ratios, Epidemiology 21 (2010) 13–15, https://doi.org/10.1097/EDE.0b013e3181c1ea43.
[26] M.J. Stensrud, J.M. Aalen, O.O. Aalen, M. Valberg, Limitations of hazard ratios in clinical trials, Eur. Heart J. 40 (2019) 1378–1383, https://doi.org/10.1093/eurheartj/ehy770.
[27] B.M. Alexander, J.D. Schoenfeld, L. Trippa, Hazards of hazard ratios – deviations from model assumptions in immunotherapy, N. Engl. J. Med. 378 (2018) 1158–1159, https://doi.org/10.1056/NEJMc1716612.
[28] R. Rahman, G. Fell, S. Ventz, A. Arfe, A.M. Vanderbeek, L. Trippa, B.M. Alexander, Deviation from the proportional hazards assumption in randomized phase 3 clinical trials in oncology: prevalence, associated factors, and implications, Clin. Cancer Res. 25 (2019) 6339–6345, https://doi.org/10.1158/1078-0432.CCR-18-3999.
[29] E. Dumas, M.J. Stensrud, How hazard ratios can mislead and why it matters in practice, Eur. J. Epidemiol. 40 (2025) 603–609, https://doi.org/10.1007/s10654-025-01250-9.
[30] M.A. Gardner, D.G. Altman, Confidence intervals rather than p values: estimation rather than hypothesis testing, Br. Med. J. 292 (1986) 746–750, https://doi.org/10.1136/bmj.292.6522.746.
[31] J. Cohen, The Earth is round ($p < 0.05$), Am. Psychol. 48 (1994) 997–1003, https://doi.org/10.1037/0003-066X.49.12.997.
[32] S. Greenland, S.J. Senn, K.J. Rothman, J.B. Carlin, C. Poole, S.N. Goodman, D. G. Altman, Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations, Eur. J. Epidemiol. 31 (2016) 337–350, https://doi.org/10.1007/s10654-016-0149-3.
[33] C. Poole, Low p-values or narrow confidence intervals: which are more durable? Epidemiology 12 (2001) 291–294, https://doi.org/10.1097/00001648-200105000-00005.
[34] F. Fidler, G.R. Loftus, Why figures with error bars should replace p values: some conceptual arguments and empirical demonstrations, Z. Psychol. 217 (2009) 27–37, https://doi.org/10.1027/0044-3409.217.1.27.
[35] R. Tomaschitz, Modeling electrical resistivity and particle fluxes with multiply broken power-law distributions, Eur. Phys. J. Plus 136 (2021) 629, https://doi.org/10.1140/epjp/s13360-021-01542-5.
[36] U. Ramer, An iterative procedure for the polygonal approximation of plane curves, Comput. Gr. Image Process. 1 (1972) 244–256, https://doi.org/10.1016/S0146-664X(72)80017-0.
[37] D.H. Douglas, T.K. Peucker, Algorithms for the reduction of the number of points required to represent a digitalized line or its caricature, Cartographica 10 (1973) 112–122, https://doi.org/10.3138/FM57-6770-U75U-7727.