Population Evolution Charts: A Fresh Look to Time-To-Event Data

Joachim Moecks [1] and Walter Koehler [2]

[1]Bioscience Club of Heidelberg, Heidelberg, Germany

[2]BaseLine- Statistical Solutions, Mannheim, Germany

Abstract: Time-to-event data analysis methods have a wide range of applications. The classical approach focuses on the survivor function, and Cox proportional hazard regression when there are covariates to assess. Population evolution charts address these data from a different angle, by concentrating on the distribution of the covariate for the developing cohort over time. Covariates that are related to the target events will show a change in distribution over time. Population evolution charts need less assumptions than the standard approaches. The present paper presents definition and interpretation of this approach for binary covariates. All conceptions and the connection to Cox-regression are detailed and illustrated by real data from clinical trials.

Published in : Statistical Computing and Graphics, Vol. 18, No 2, Dec 2007, p. 12-19

http://stat-computing.org/newsletter/

Acknowledgement: The work on Population Evolution Charts started in 1995 while working for Boehringer Mannheim / Clinical Development and was later continued within Roche Pharma /Global Development. In this phase, we greatly acknowledge contributions of Martin Scott, Jörg Maurer, Michael Budde and Sam Givens. The recent completion of the work was stimulated by fruitful PEC applications in Roche's biomarker development, while the authors supported Roche Basel as consultants in biomarker statistics.(This acknowledgement was unfortunately not published in Statistical Computing and Graphics).

Correspondence to: J. Moecks Joachim@Moecks.net

Introduction:

Time-to-event data (TTE data) is generated in many fields of science and engineering. The methodology for analyzing these data originates in the analysis of life table records – hence calling it 'survival analysis', whereas in more technically-minded environments the same thing was called 'analysis of failure-time-data'. We prefer the more neutral term 'event', however, the methodology cannot deny its origin by considering quantities called 'survivor function' or 'hazard'. However named, this represents an emerging field of applied statistics enjoying a wide range of applications from social science, economy, quality and reliability testing to medicine. Medicine represents probably the largest area of application, where statistical TTE-analysis became a key technology of data analysis, whether in epidemiology or in clinical trials. The methods considered in this article were developed and motivated from clinical trials and will be illustrated by means of those data, but may be applied in any other field of TTE-analysis.

In many clinical trials the success of therapeutic strategies is described by time variables, e.g. "time to some clinical event" or "time to therapy response/cure". In the field of serious diseases e.g. cancer, many studies consider clinical endpoints like "time to disease progression", "time to death (overall survival, OS)", or in combination, e.g. "time to progression or death (progression free survival, PFS) ".The success of a therapy is evaluated by investigating the distribution of the random variable T (time-to-event), through the survivor function $S(t) = 1 - F(t) = P(T > t)$ giving the probability that the event occurs after t.

The survivor function as such is in most cases not of primary interest, but the relation of survival to covariates, i.e. to compare the risk, e.g. for two treatment arms, for the two sexes, for a pre-existing disease (yes or no), or for two groups defined by the cutoff of a metric covariate, say. (There could also be metric covariates – but we will concentrate on binary cases in this article).

Population Evolution Charts (PECs) conceive the follow-up of patients as a selection process relative to the covariate distribution: suppose at study start (baseline) there are p% males in the cohort. Whenever an event occurs a patient leaves the cohort i.e. is selected. If the selection of males proceeds with a same base probability of p% over the whole study duration, then the selecting events are indifferent (unrelated) with respect to the covariate. However, if systematically more or less males than p% were selected then a relation of events and covariate can be assumed. This simple idea behind the PEC may be written a little more formally:
Consider a binary baseline covariate X with observations $x_j \in \{0,1\}$ where $j \in G(0)$, the studied cohort at time *t=0* (baseline, start of the study). The initial cohort $G(0)$ reducing it to some $G(t)$ at study time *t*. In order to investigate whether something changed with the cohort with respect to the covariate X simply compare the fraction of those with X=1 at baseline $\bar{x}(0) = \frac{1}{|G(0)|} \sum_{j \in G(0)} x_j$ with $\bar{x}(t) = \frac{1}{|G(t)|} \sum_{j \in G(t)} x_j$, the fraction of those with $X = 1$ for the remainder cohort at time *t*. A plot of $\bar{x}(t)$ represents the basic Population Evolution Chart (PEC). If the events are not associated with the covariate, we expect a constant course of the PEC , i.e. the selection process selects indifferently out of the two groups and thus will not change the composition. Systematic deviations from constancy speak for an association of X with the selection process, e.g. if $\bar{x}(t)$ tends downwards then those with $X = 1$ are at a higher risk respectively earlier risk to perceive an event. It is an attractive feature of PECs that the association and its time pattern can be studied without any difficult assumptions.

Two examples of PECs from real clinical data are given in Figure 1. The PEC for male percentage indicates that sex is clearly related to the considered events: after about 120 days of treatment the

percentage of males in the study population dropped down from 70% to 60%, the development shows a clear monotonic trend (interpretations will be studied below). The green trace demonstrate how a independence will show up in a PEC, here showing this for a random covariate, which by construction is not associated with the event times.
In the PEC example, we have tacitly ignored the problem of censoring in time - to- event data, we will come back to this in the estimation section.

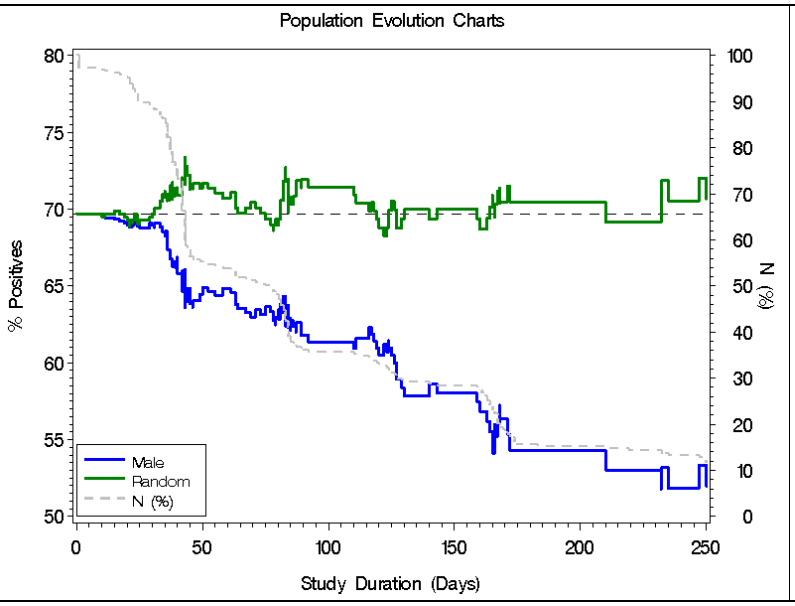

Figure 1: Population Evolution Chart Examples. Blue trace: The percentage of males in development over treatment time. Green trace: binary random covariate with no relation to events (same event times as for blue trace)

Setting and Notation

To avoid any technical difficulties, we will assume in the following that all studied quantities are mathematically well behaved. Here are in brief the main 'coordinates' of the standard time-to-event terminology: The analysis is not based on the distribution function, i.e. $F(t) = P(T \le t)$ but rather makes use of the survivor function $S(t) = 1 - F(t) = P(T > t)$, giving the probability that the event occurs after t. The ubiquitous quantity in TTE-analysis is the hazard function $h(t)$, describing the risk for an immediate event at *t* conditional on survival up to *t*. It can be obtained from the survivor function together with the integrated hazard $H(t)$ as follows:

$$H(t) = -\log(S(t))$$

$$h(t) = {}^{d}\!/_{dt} H(t) = -{}^{d}\!/_{dt} \log(S(t)) = \frac{f(t)}{S(t)},$$

where $f(t)$ denotes the density of T (Kalbfleisch & Prentice, 1980, p. 6; Lee 1992, chapter 2).

The standard approach for studying the influence of covariates is to consider the conditional distribution of T i.e. $P(T > t \mid X = 1)$ and by fitting some (semi-)parametric model (e.g. Cox proportional hazard model, Kalbfleisch & Prentice, chapters 4, 5).

In contrast, population evolution charts emphasize the covariate and take the opposite view to the TTE-data: The PEC can be straightforwardly defined as a sequence of conditional probabilities:

(1) $$\Psi_X(t) := P(X = 1 \mid T > t)$$

The population evolution chart thus represents the probability for $X = 1$ for the survivors up to t.

The PEC was introduced as describing a selection process – this is underlined by the following derivation. Consider the distribution of the covariate *X* at time 0 (baseline): $P(X = 1)$. The simple identity

$$P(X = 1) = P(X = 1; T > t) + P(X = 1; T \le t) \ ,$$

leads by elementary probability algebra to

$$P(X = 1) = P(X = 1 \mid T > t) \cdot P(T > t) + P(X = 1 \mid T \le t) \cdot P(T \le t) \ ,$$

By putting $S(t) = P(T > t)$, the overall survivor function, we thus get:

$$\Psi_X(t) = P(X = 1 \mid T > t) = \frac{1}{S(t)} \left[P(X = 1) - (1 - S(t)) \cdot P(X = 1 \mid T \le t) \right] \tag{2}$$

This underlines that the events can be conceived to select out of the baseline distribution $P(X = 1)$, leading to a time dependent composition of the population.
A further representation of the PEC can be directly found from its definition as conditional probability:

$$\Psi_X(t) = P(X = 1 \mid T > t) = \frac{P(T > t \mid X = 1)}{P(T > t)} P(X = 1) = \frac{S_1(t)}{S(t)} P(X = 1) \tag{3}$$

Estimation

The well known complication of time-to-event data is the fact that not all events are actually observed, but the observation may terminated at *t* before an event occurred (censoring), e.g. due to fixed administrative date for study end. For censored observations it is only known that the event did not occur up to *t*. The well known answer to this issue are special estimates of the survivor function, like the celebrated Kaplan-Meier graph.

In the case of no censoring, there is no estimation problem and we simply proceed as indicated in the Introduction. However, if there is some censoring, then the cohort $G(t)$ is not complete and could miss some of those censored before *t*, for it is unknown whether they would have survived *t*. Censoring comes usually with assumptions e.g. that it is random or independent of the time variable *T*. In most cases it is also customary to assume that censoring is not dependent on the covariates, e.g. that any demographic property of the patient or pre-existing co-morbidities would influence the censoring. In clinical trials, we actually face two types of censoring, the pure administrative date of study end or time of interim analysis and censoring that occurs during the conduct of the study, patients drop out voluntarily or for unknown reasons, e.g. due insufficient effect of the control treatment (placebo) or side effects of the test treatment. A complete independence of these censored observations from the events or from covariates is usually assumed, but may be overoptimistic.

In order to address different assumptions regarding we will conceptually introduce censoring into the setting:

Let *C* denote the r.v. of the censoring process. The r.v. pertaining to all actually observed times (censoring and events) will be denoted by *A*. Note that $A = \min(T, C)$. If censoring and events are independent, we get for the total process multiplicative survivor functions:

$$P(A > t) = P(T > t) \cdot P(C > t)$$

This property carries directly over to the Kaplan-Meier (KM) estimates of the involved survivor functions:

$$\hat{P}(A>t)=\hat{P}(T>t)\cdot\hat{P}(C>t)$$

Note that $\hat{S}(t)=\hat{P}(T>t)$ is the usual KM-estimate of the survivor function, while for calculating $\hat{P}(C>t)$, the role of events and censored observations is to be interchanged.

Let us now consider the PEC for the total process pertaining to *A* according to representation (3):

$$P(X=1\mid A>t)=\frac{P(A>t\mid X=1)}{P(A>t)}P(X=1)$$

By independence of *T*, *C* we get:

$$P(X=1\mid A>t)=\frac{P(T>t\mid X=1)}{P(T>t)}\cdot\frac{P(C>t\mid X=1)}{P(C>t)}\cdot P(X=1)$$

The assumption of independence of covariate *X* and censoring *C* leads to $P(C>t\mid X=1)=P(C>t)$ and hence we get:

$$P(X=1\mid A>t)=\frac{P(T>t\mid X=1)}{P(T>t)}\cdot P(X=1)=P(X=1\mid T>t)$$

In this case the PEC could be estimated based on the total process *A*. The simple estimate presented in the introduction also refers to the total process and thus could serve as first try in case of independent censoring:

(4) $$\hat{\Psi}_{\mathrm{I}}(t)=\tfrac{1}{|G(t)|}\sum_{j\in G(t)}x_j$$

A further way of estimation is offered by the representation from (3): Let $\bar{x}(0)=\tfrac{1}{|G(0)|}\sum_{j\in G(0)}x_j$ denote the baseline estimate of $P(X=1)=\Psi_X(0)$ and let $\hat{S}(t)$, $\hat{S}_1(t)$ denote the KM- estimate of $P(T>t)$, $P(T>t\mid X=1)$, respectively, which leads to:

(5) $$\hat{\Psi}_{\mathrm{II}}(t)=\frac{\hat{S}_1(t)}{\hat{S}(t)}\cdot\bar{x}(0)$$

In comparing the two estimators, we observe that (5) requires the least assumptions, since there is no need to assume that the covariate is independent of censoring. A further advantage of (5) is that we may investigate the censoring process by defining a PEC for censoring:

(6) $$\hat{\Psi}_{\mathrm{C}}(t)=\frac{\hat{P}(C>t\mid X=1)}{\hat{P}(C>t)}\cdot\bar{x}(0)\,.$$

The estimates derive from KM-estimates with censoring and events interchanged. Since (4) refers to the total process, censoring and events cannot be separately studied.

The two estimators of the PEC are illustrated in figure 2, the example shows the results for a biomarker, where positivity indicates a worse medical prognosis. It is seen that there is a strong decline in the PEC for events for the estimator II, which means that indeed positivity for this

parameter leads to more events. At the same time a different development is seen for the censor CPEC. It shows a slight trend to the opposite direction, which means that those censored are with higher probability negatives with regard to the biomarker. The trace for the estimator I shows that we have to deal with variance for larger t, and that in particular for larger t, the estimates are affected by the censored events.

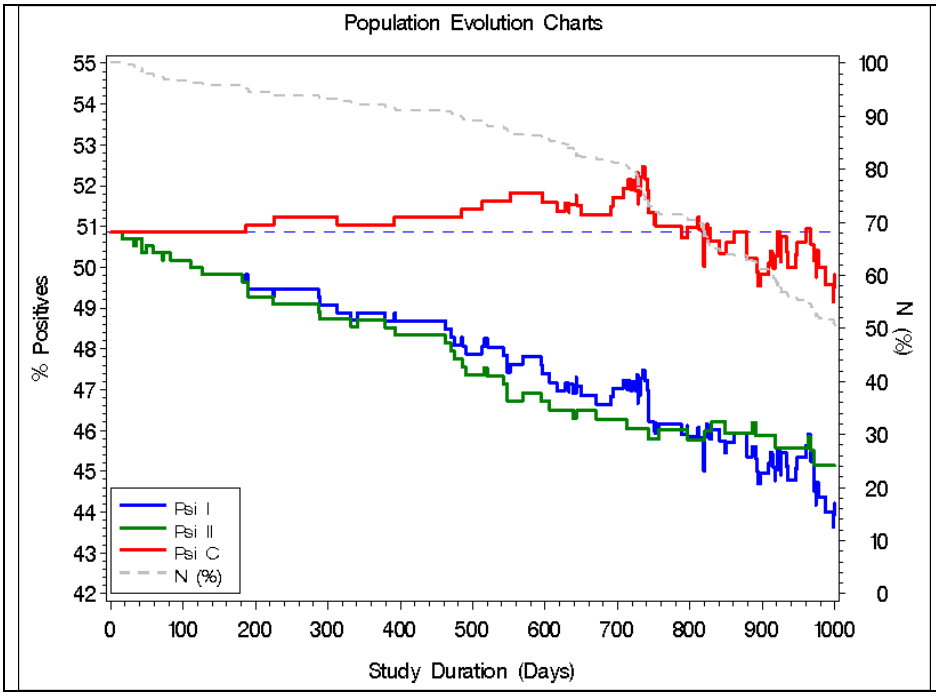

Figure 2: Estimates of the Population Evolution Chart. The real data example of a baseline biomarker. Patients with biomarker positivity are predominantly selected by the considered events. Censoring events show a weak trend in the opposite direction.

Interpretation of PECs

The basic goal of a PEC analysis is to investigate dependencies in a nonparametric way. The overall null-hypothesis is the constancy of the PEC, i.e. that there is no relationship of events and covariate. Formally we have:

$$\text{H0:} \quad \Psi_X(t) := P(X = 1 \mid T > t) = \text{const.}$$

The constant is the value of the PEC at t=0, i.e. $P(X=1) = \Psi_X(0)$. From the representation (3) we get the equivalent hypothesis $S(t) = S_1(t)$ which is also equivalent to $S_1(t) = S_0(t)$, where $S_0(t) = P(T > t \mid X = 0)$ denotes the survivor function for $X = 0$. Thus for testing the overall constancy of a PEC, we are facing a standard problem in TTE-methodology, namely to compare two survivor functions, which can be well done non-parametrically e.g. by the logrank or Wilcoxon test.

In order to look a bit deeper into the guts of the PEC approach, we note that $S(t) = S_1(t) \cdot P(X=1) + S_0(t) \cdot P(X=0)$. By taking the derivatives of the log of this relation,

$$h(t) = -{}^{d}\!/_{dt} \log\{S_1(t) \cdot P(X=1) + S_0(t) \cdot P(X=0)\} = S(t)^{-1}\{f_1(t) \cdot P(X=1) + f_0(t) \cdot P(X=0)\}$$

some algebraic arrangement leads to:

(7) $$h(t) = h_1(t) \cdot \Psi_X(t) + h_0(t) \cdot (1 - \Psi_X(t))$$

This provides a useful relation for the overall hazard linking it to the PEC and pertaining subgroup hazards.
Further, we may write the PEC as a function of underlying hazards for $h_1(t) - h_0(t) \neq 0$:

(8) $$\Psi_X(t) = \frac{h(t) - h_0(t)}{h_1(t) - h_0(t)}$$

The interesting case for a PEC is the deviation from constancy – and here is the question what is the meaning of the speed of deviation from constancy. Taking log in the representation (3), and taking subsequent derivatives, leads to:

$$\log(\Psi_X(t)) = \log(P(X = 1 \mid T > t)) = \log S_1(t) - \log S(t) + \log P(X = 1)$$

(9) $$\frac{\Psi_X'(t)}{\Psi_X(t)} = h(t) - h_1(t)$$

The dynamics in the change of the PEC is determined by a hazard difference. This relation can be transformed to involve the subgroup hazards $h_1(t)$ and $h_0(t)$:

(10) $$\Psi_X'(t) = (h_0(t) - h_1(t)) \cdot \Psi_X(t) \cdot (1 - \Psi_X(t))$$

Since the PECs are positive functions, only the hazards determine whether the PEC is going up or down: whenever $h_1(t)$ exeeds $h_0(t)$ the PEC is going down and vice versa, as local instantaneous property. PECs will get a monotonic course when one of the hazards is dominating the other for all considered times. The PECs shown in figures 1 and 2 very well follow this pattern.

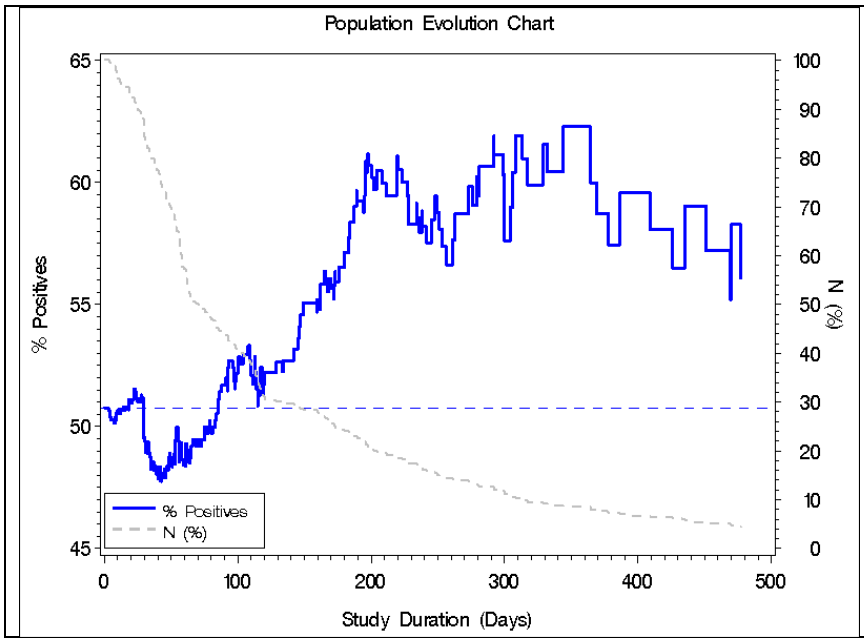

Figure 3: A non-monotonic Population Evolution Chart (real data). A turn point is visible approximately after 40 days.

The PEC shows in figure 3 shows a different pattern: after an initial phase with very little difference in hazards, there was a clear drop, when the positives had more events. After about 40

days the picture changed and the negatives suffered from more events. Such a pattern fits to the idea that a positive status is associated with a bad prognosis – however, after a lag time of about 40 days, the therapy is able to change the picture creating specifically for positives a benefit from this therapy, in a clearly higher degree than for the negatives. Later, after approx. 200 days, the PEC becomes more or less constant, which could mean that those of the negatives, who have survived up to this point, do no longer suffer from their negative status.

PECs and Cox' proportional hazard model

The semi-parametric Cox proportional hazard model represents the gold standard for dealing with covariates in the TTE framework. It represents probably one of the most often used statistical approaches at all – at least in the medical context. The crucial assumption of this model is the proportionality of the hazard functions of the two groups ($X=1$ and $X=0$), i.e. their hazard ratio is independent of time. We will only consider the simple case of one binary covariate.

The gist of the difference between Cox model and PEC approach is easily stated: the Cox model looks at $P(T>t \mid X=1)$, while the PEC features $P(X=1 \mid T>t)$. PECs study the time development of the covariate distribution, while in the Cox-model the covariate influence is independent of time. The influence of the covariate is parametrically quantified by modelling $P(T>t \mid X=1)$ with proportional underlying hazards, without restriction of the functional form.

It is illustrative to investigate the form of a PEC, when the assumptions of a Cox model hold, i.e $h_1(t)=\lambda \cdot h_0(t)$. Then we get from (10):

(11) $$\Psi'_X(t)=(1-\lambda)\cdot h_0(t)\cdot \Psi_X(t)\cdot(1-\Psi_X(t)).$$

This means that a PEC from a valid Cox model would always be monotonic, for $\lambda>1$ decreasing and for $\lambda<1$ increasing. This comes as a necessary condition for the validity of the Cox model, but monotonicity of the PEC is not sufficient: by (10) , we may have $h_1(t)>h_0(t)$ for all t and thus a monotonic PEC, but $h_1(t)=\lambda(t)\cdot h_0(t)$ with a non-constant $\lambda(t)$ which violates the proportionality assumption.

In the case of figure 3 it is very doubtful, whether a simple Cox model can render reasonable results. Maybe a piecewise approach can be tried by considering the data between day 40 and 200 to be fitted by a Cox approach.

The PECs in figures 1 and 2 show a perfect linear decline, what can we read out of this, when we assume that a Cox model is valid ? To this end, suppose that the PEC is really linear:

$$\Psi_X(t)=\alpha\cdot t+\pi_1 \quad ,$$

where the intercept is $\pi_1=\Psi_X(0)=P(X=1)$. Using (11) and putting $\pi_0=1-\pi_1$ we find:

(12) $$h_0(t)=\frac{1}{1-\lambda}\cdot\frac{\alpha}{(\alpha\cdot t+\pi_1)\cdot(-\alpha\cdot t+\pi_0)}$$

Thus, a simple parameterization of the PEC leads to an explicit functional form for Cox' baseline hazard. In order to provide a real example, the involved parameters were estimated from the data

of the figure2 example. The results were as follows: $\pi_1 = 50.86\%$ (obtained from PEC), $\alpha = 4.88 \frac{\%}{1000d}$ (obtained by linear regression), and $\lambda = 3.24$ (obtained from Cox regression). The resulting hazard estimates according to (12) are displayed in figure 4. The hazards are almost constant, the small embedded picture shows the baseline hazard with a magnified scale – demonstrating the slight increase of the hazard over the time span.

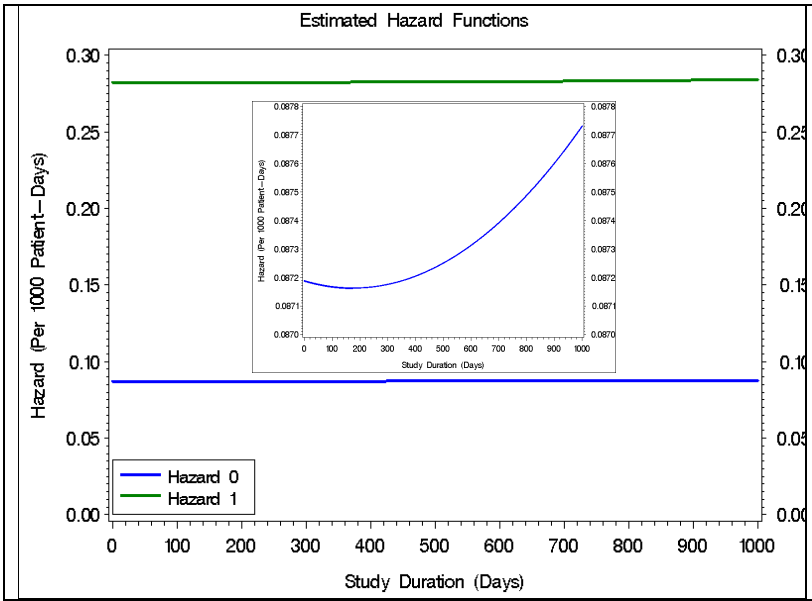

Figure 4: Estimated hazard functions according to (12) based on the data displayed in figure 2. Blue trace: baseline hazard; green trace: hazard for biomarker positives. Embedded graph: baseline hazard in magnified scale.

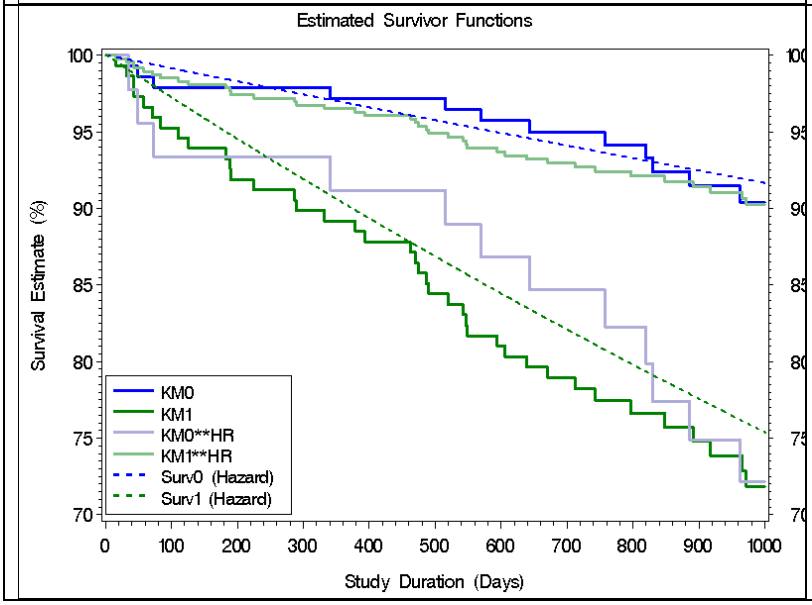

Figure 5: Estimated survivor functions based on the data displayed in figure 2. Blue trace(KM0): Kaplan – Meier estimate for group of biomarker negatives. Green trace(KM1): Kaplan – Meier estimate for group of biomarker positives. Light blue($KM0^{3.24}$): estimated survival of biomarker positives according to the Cox estimate of hazard ratio 3.24. Light green($KM1^{1/3.24}$): estimated survival of biomarker negatives according to the Cox estimate of hazard ratio 3.24. Dotted lines: estimated survivor functions according to the hazard estimates of figure 4.

Figure 5 displays survivor functions for the subgroups estimated by various approaches. The fit to the Cox model is fair (as seen from the estimated survivor functions based on $\lambda = 3.24$ (light colors) in comparison to the standard Kaplan Meier subgroup estimate. The almost exponential survivor functions estimated according to (12) provide a good compromise in these data.

Discussion

The basic idea of PECs was to consider TTE-data as a selection process and to look for the time development of the observed cohort. This leads to quite interesting graphical displays which allow

for an alternative view to the TTE-data and covariates, in particular by detailing the time pattern of the covariate influence.

The idea of PECs evolved in the year 1996, when the cardio-vascular safety of recombinant erythropoietin (EPO) for dialysis patients was analyzed in a large study pool with initially 3111 patients of 22 clinical trials with variable follow-up for up to 4 years. The estimated course of the hazard for cardio-vascular-death showed a surprising decline over the follow-up period (see Moecks et al., 1997). Could this represent a long term benefit of the anti-anemic therapy? Was the case mix of the population starting the first year (n=3111) comparable to those starting the fourth year (n=101) ? CV-death could select out those patients who are anyway prone to this risk, leaving only comparably healthy patients at later stages. The study pool contained studies with differing follow-up and different inclusion/exclusion criteria, with an unclear effect on the remainder cohort. Therefore the PEC addressed the total selection process in particular for baseline covariates which implied a cardiovascular risk. The PECs revealed that there was no decrease in percentage of patients with pre-existing CV-disease or diabetes, rather these percentages even raised. The case mix of patients regarding CV-risk was well comparable between first and fourth year and in between, thus supporting that the observed hazard decrease was a long term treatment benefit (Moecks, 2000).

The second practical instance where PECs proved useful, came up in the context of bisphosphonate therapy for bone morbidity in cancer patients. The endpoints were bone events (bone pain, pending fractures etc) which should be reduced by therapy compared to control. Here premature dropouts presented with issues, since those with advanced morbidity under the (inefficient) control therapy, showed a higher dropout rate, with a downward bias in reported bone events. PECs revealed that dropout was selective, i.e. tended to select out patients with advanced morbidity. Moreover, in some studies, the dropout effect was different for active and control, showing that more morbid patients dropped out at a higher rate under control (Moecks et al, 2002). This underlines efficacy but gives standard approaches a hard time to show efficacy.

A important merit of the PEC approach in the context of clinical trials is the possibility to use this chart for censor events: suppose that we find a covariate which clearly is associated with the target events, and in addition a PEC for censoring reveals a dependency as well. Then the censoring could violate the customary assumptions and exert a bias in the one or other direction, e.g. as indicated in the bone event example.

The derivations of this article show that the PEC serves as a basic descriptor in the TTE-methodology. For instance equation (7) shows that the overall hazard and the subgroup hazards are linked through the PEC – a quite fundamental relation, linking these intuitive descriptors of the risk development.

The present article only dealt with the simplest case of a binary covariate. Displays similar to PEC's can also be defined and used for metric covariates. Further a couple of more testing options exist in order to get a probabilistic evaluation of the PEC-course. The focus of a PEC is on one single covariate, and the conception does not readily generalize to the multiple situation. It is however possible to treat one binary covariate (e.g. treatment arms) as subgrouping variable and to compare PECs for a further covariate in one graphical display (e.g. a PEC for each treatment arm separately).

In the meantime PECs had been applied in many data analyses, in particular in the context of diagnostic markers and biomarkers in oncology, providing many fruitful insights. A further development of this methodology appears promising for the TTE-field.

Computations of this paper were based on SAS-code; a macro for the PEC-estimate $\hat{\Psi}_{II}(t)$ can be obtained from the authors upon request.

References

Kalbfleisch J.D., Prentice R.L.(1980): The Statistical Analysis of Failure Time Data. John Wiley &Sons: New York

Lee E.T.(1992): Statistical Methods for Survival Data Analysis – Second edition. John Wiley &Sons: New York

Moecks J., Franke W., Ehmer B., Quarder O. (1997): Analysis of Safety Database for Long-Term Epoetin-β Treatment: A Meta-Analysis Covering 3697 Patients. In: Koch & Stein (editors) "Pathogenetic and Therapeutic Aspects of Chronic Renal Failure". Marcel Dekker: New York, 163-179.

Moecks J. (2000): Cardiovascular mortality in haemodialysis patients treated with erythropoietin – a retrospective study; Nephron, 86; 455-462

Moecks J., Köhler W., Scott M., Maurer J., Budde M., Givens S.(2002): Dealing with Selective Dropout in Clinical Trials. Pharmaceutical Statistics, Vol 1: 119-130