Collection of Biostatistics Research Archive
COBRA Preprint Series

Year 2010 *Paper* 70

Improving statistical analysis of prospective clinical trials in stem cell transplantation. An inventory of new approaches in survival analysis

Aurelien Latouche*

*University of Versailles, aurelien.latouche@uvsq.fr

Improving statistical analysis of prospective clinical trials in stem cell transplantation. An inventory of new approaches in survival analysis

Aurelien Latouche

Abstract

The CLINT project is an European Union funded project, run as a specific support action, under the sixth framework programme. It is a 2 year project aimed at supporting the European Group for Blood and Marrow Transplantation (EBMT) to develop its infrastructure for the conduct of trans-European clinical trials in accordance with the EU Clinical Trials Directive, and to facilitate International prospective clinical trials in stem cell transplantation. The initial task is to create an inventory of the existing biostatistical literature on new approaches to survival analyses that are not currently widely utilised. The estimation of survival endpoints is introduced, with an emphasis on recent developments which complements standard analysis. The issues raised are new regression models that allow the estimation of time dependent effect for cause specific hazard, cumulative incidence and more generally mean response. New development in multi state model, with notably, recent regression models that assess the influence of covariates directly on transition probabilities are detailed. Some recent test for comparing cumulative incidence function across treatment arm are introduced. The estimation of centre effect in multi centric studies is also documented. Sample size calculation in the presence of competing risks are then presented. We close with the inventory of available packages and macro in R that implement the previous survival models.

Improving statistical analysis of prospective clinical trials in stem cell transplantation. An inventory of new approaches in survival analysis

A. Latouche
INSERM U1018, Biostatistics Team
Université de Versailles, France

Abstract

The CLINT [1] is an European union-funded project, run as a specific support action, under the Sixth Framework Programme. It is a 2 year project aimed at supporting the European Group for Blood and Marrow Transplantation (EBMT) to develop its infrastructure for the conduct of trans-european clinical trials in accordance with the european union clinical trials directive, and to facilitate international prospective clinical trials in stem cell transplantation. The initial task was to create an inventory of the existing biostatistical literature on new approaches to survival analyses that were not currently widely utilised.

The estimation of survival endpoints is introduced, with an emphasis on recent developments which complements standard analysis. The issues raised are new regression models that allow the estimation of time dependent effect for cause specific hazard, cumulative incidence and more generally mean response. New development in multi state model, with notably, recent regression models that assess the influence of covariates directly on transition probabilities are detailed. Some recent test for comparing cumulative incidence function across treatment arm are introduced. The estimation of centre effect in multi centric studies is also documented. Sample size calculation in the presence of competing risks are then presented. We close with the inventory of available packages and *macro* in R that implement the previous survival models.

Keywords: clinical trial; competing risks; multistate model; centre effect; sample size

1 Introduction

Patients who undergo a hematopoietic graft, can encounter several events post transplant: namely engraftment, graft–versus–host–disease, relapse, non-relapse death, progression. To assess the effect of a treatment on such outcome, some specific survival model are needed. The Cox proprotional hazards dominates the survival analysis for years, notably because of the ease of the interpretation. The use of this model is perfectly detailed in the classic book of Therneau and Grambsch [1]. The topics covered are : residuals analysis to test the proportional hazards assumption, the functional form of the covariate or influence of individuals, time–dependent effect/coefficient (time–varying effects), correlated observations such as repeated measures and frailty or random effects. Other textbooks include Klein and Moeschberger [2] (with an emphasis on hematology case studies), Kalbfleish and

[1] Establishment of infrastructure to support international prospective clinical trials in stem cell transplantation (CLINT)

Prentice [3], Hosmer and Lemeshow [4] and the revised edition Hosmer et al. [5] as well as Collett [6], Kleinbaum and Klein [7].

The hematology field is a very inspiring when it comes to statistical developments especially in survival analysis. A first look at the litterature confirms this interest notably through methodological notes or review that populate medical reviews [8, 9, 10, 11]. Theses notes and articles focuse mostly on comparison of regression for hazard rates and cumulative incidence functions [12, 13]. In this inventory we will consider, alternative modelling strategies that complements the traditional proportional hazards model as well as estimations presented in the book of Therneau and Grambsch [1]. Recent books covering the topics are: Handbook of Statistics 23 [14], Dynamic Regression Models for Survival Data [15].

Before, investigating what were the advances since 2000 in survival analysis. It is of interest to list what are up tu now the major tool at–hands. Major advances in survival analysis are the Survival R–package by T. Therneau [16], Multistate modelling, Tests for comparison of cumulative incidence functions [17], Regression model for the cumulative incidence and Fine–Gray model, Additive hazard model (Aalen, Scheike) [15]. All theses points will be exemplified in the sequel.

In the main hematological reviews there are very comprehensive recommendation on the respective merits of up–to–date methods. This is mostly due to JP Klein (Medical College of Wisconsin) and colleagues that disseminate appropriate methodologies in the stem cell transplantation field [8, 9]. For exemple in a review paper Kim [11] introduced the pseudo–value estimation method for regressing the cumulative incidence functions. This method is new but is already made available (in principle) to applied statisticians. Another striking, exemple is the dissemination of the Fine–Gray model for the subdistribution of a competing risk. This is mostly due to the availability of a R–package. Indeed, the lack of statistical software that implements novel methodolodgies leads to underuse models. In that respect, in this inventory, we focused on model with ready–to–use software or routines.

In the first part , we recall standard notation and statistical models. Next we introduce prognostic factor analysis with the regression modelling and hypothesis tests. We close with a synthesis of statistical softwares and add–on package.

2 Statistical Models

In this section, we introduce the major statistical models when the interest is the analysis of time–to–event failure.

2.1 Survival model

The standard survival model focuses on a single endpoint. Recent developpments are numerous, notably we identified, alternative methods (or tests) for the comparison of survival curves.

Usually, comparison of survival curves among randomization arms are performed at a fixed time–point. Klein et al. [18], investigated the performance of naive test (difference between the two survival curves).

Logan et al. [19] focused on crossing survival curves (that contradict the PH assumptions). A number of methods for comparing two survival curves after a prespecified time point. This situation may be of interest when the survival curves are expected to cross, so that we are only interested in late difference. Another, recent developpments is the study of alternative endpoint, such as the Progression Free Survival (PFS) [20].

2.2 Multistate models

Since the papers of Klein et al. [21], Keiding et al. [22] the multi–state approach, is becoming more popular but remain solely in hematopoeitic stem cell transplantation (HSCT). The use of multistate in HSCT is not particlularly new [21, 23]. One possible reason is that a multi-state model regression analysis typically involves the modelling of each transition intensity separately. Each probability of interest, namely the probability that a subject will be in a given state at some time, is a complex nonlinear function of the intensity regression coefficients. Thus, interpertation in terms of probability is quite complicated (even if depict the patient more closely) and the interpretation of hypothetical predictions from multi-state models in HSCT have to be avoided. An interesting exemple of the versatility of the multi-state model is the Current Leukemia Free Survival. In this exemple, the patient move between 9 states [24, 25, 26].

There exists an extensive literature on multi-state model. Main contributions include books by Andersen et al. [27] and Hougaard [28]. Recent reviews on this topic may be found in Hougaard [29], Andersen and Keiding [30]. An issue of the Journal Statistical Methods in Medical Research, entirely devoted to these models, was published in 2002. Despite its potentialities, multi-state modelling is not used by practitioners as frequently as other survival analysis techniques. Lack of knowledge of the available software as well as misunderstanding of what multi-state modellings advantages rely on (compared to the simple Cox model), are probably responsible for this lack of popularity.

The paper of Andersen et al. [31] entitled *Competing risks as a multistate model*, gave a fresh and unified view about multi–state model and competing risks. Recent developpement of Scheike and Zhang [32] that suggested a direct modelling of regression effects for transition probabilities should bring this framework up–front.

R–script of the tutorial from Putter et al. [33] can be found at `http://www.msbi.nl/multistate`. For a comprehensive review, we suggest the work of Meira–Machado and `tdc.msm` script [34]. More recently a R–package `mvna` provides plots and estimates of the cumulative hazards as a function of time for all the transitions specified by the user.

2.3 Competing risks model

For simplicity and tractability we will consider 2 competing events to introduce fundamental quantities.. In HSCT setting, this will usually be relapse and death in remission *aka* non relapse mortality.

The observed data typically consist in an observation time $\tilde{T}$ which is the minimum of a failure time T and a censoring time C and a status indicator ε. $\varepsilon = 0$ if the observation is censored ($C < T$). If $T > C$, then ε denotes the observed cause of failure with $\varepsilon = 1$ for the event of interest and $\varepsilon = 2$ for the other competing event. Most common analyses focus on comparing the *cause–specific* hazard under the control and the experimental treatment [35], where the cause–specific hazard of failure from cause 1 in treatment arm E (resp C for control) is defined as:

$$\lambda_{1E}(t) = dF_{1E}(t)/S_E(t)$$

with F_{1E} is the cumulative incidence function of failure from the cause of interest, *i.e.* $F_{1E}(t) = \Pr(T \leq t, \epsilon = 1)$ and $S_E(t) = 1 - (F_{1E}(t) + F_{2E}(t))$ is the event free survival function. In such a case, comparisons of cause–specific hazards between groups are performed against proportional hazards alternatives, using a Cox model. The other strategy consists in comparing the corresponding event probabilities $F_{1E}(t)$ and $F_{1C}(t)$, either directly the Gray's test [17] or using a Cox-like model for the associated hazard $\alpha_{1E}(t) = dF_{1E}(t)/(1 - F_{1E}(t))$, referred as the subdistribution hazard [13].

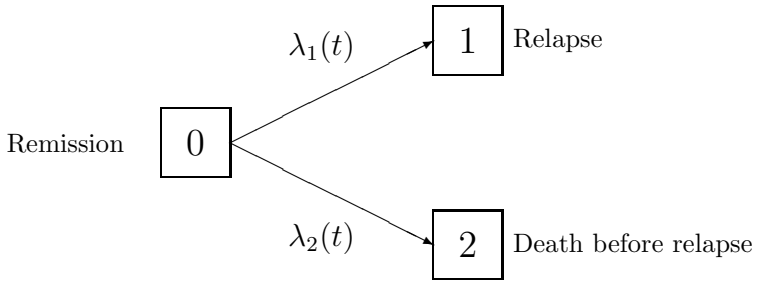

Figure 1: Competing risks model with cause-specific hazard for relapse $\lambda_1(t)$ and cause-specific hazard for death $\lambda_2(t)$.

The subdistribution hazard is directly related to the cumulative incidence function while the relation between cause–specific hazard and cumulative incidence involves the cause–specific hazard of the competing event. Another key remark is that these two models cannot hold simulatneously, *i.e* proportional cause–specific hazards imply non–proportional subdistribution hazards.

It seems now established that these two models should be used simultaneously to fully depict the complex course of the patients. A detailed discussion of the relative merits of both approaches and their interpretation can be found in the paper of Beyersmann and Schumacher [36]. Tutorial on competing risks analysis are provided by Putter et al. [33] (with EBMT data) and in other field such time–to–seroconversion collaborative group such as CASCADE [37] provide guidelines for analysis of competing risks data.

To estimate the treatment benefit, it is recommended to adjust treatment comparison for potential confounders, based on regression models. For competing risks data, two main approaches have been used, either the Cox model [12] or Therneau and Grambsch [1, Chapter 8.4], or the recently proposed Fine and Gray model [13] and Martinussen and Scheike [15, Chapter 10]. Despite its rather recent origin, the Fine–Gray model has been quickly put to use in applications such as neutrophils recovery after bone marrow transplantation [38], infectious complications after blood stem-cell transplantation [39].

The models take subtlety different approaches to competing risks data, and it is important to understand this for proper interpreting these respective results. Both the Cox and the Fine–Gray model analyze data from a competing risks setting as displayed in Figure 1. We observe a so-called failure time T between start of remission and relapse/death, whatever comes first. One can think of time T as the time spent in the remission state 0 until moving into one of the competing risk states 1 (relapse) or 2 (death without prior relapse). Attached to these competing risks are cause-specific hazards $\lambda_1(t)$ and $\lambda_2(t)$; these can be thought of as 'instantaneous forces' that draw an individual towards the respective competing risks. More precisely, $\lambda_1(t)$ multiplied by a very, very small time interval is the probability of relapsing within this small time interval under the condition that the individual has still been relapse-free at the beginning of this time interval.

Hazards are, in fact, a very elusive concept [40], but analysis and interpretation is straightforward in usual survival analysis. A usual Cox model would look at the all-cause hazard $\lambda(t) = \lambda_1(t) + \lambda_2(t)$, which has a one-to-one correspondence to the distribution of the failure time T through

$$P(T \leq t) = 1 - \exp\left(-\int_0^t \lambda(u)\,\mathrm{d}u\right),$$

i.e. the proportion of patients experiencing death or relapse (whatever comes first), as time progresses. Due to this one-to-one correspondence, a decreasing treatment effect found in a Cox model means a

decrease in this proportion, and an increasing effect entails an increase in this proportion.

However, things become surprisingly difficult with competing risks. We may still fit Cox models, but as is apparent from Figure 1 we will need to fit two Cox models, one for each cause-specific hazard, see, e.g., [1, p. 177]. The interpretation of these results then becomes involved, because the CIF for relapse, say, depends on both cause-specific hazards, and it does so in a rather complicated way [41]. In fact, we have for the CIF of relapse

$$\mathrm{CIF}_1(t) = P(T \leq t, \text{Relapse at } t) = \int_0^t \exp\left(-\int_0^u \lambda_1(v) + \lambda_2(v)\,\mathrm{d}v\right) \cdot \lambda_1(u)\,\mathrm{d}u$$

These difficulties have led to the Fine–Gray model [13], with the aim of doing a Cox-type analysis for a quantitity *which reestablishes the one-to-one correspondence to the CIF of relapse*.

This quantitity has come to be known as the subdistribution hazard for relapse, and we write $\lambda^{\mathrm{FG}}(t)$ for it. The aim is to reestablish

$$\mathrm{CIF}_1(t) = P(T \leq t, \text{Relapse at } T) = 1 - \exp\left(-\int_0^t \lambda^{\mathrm{FG}}(u)\,\mathrm{d}u\right)$$

for the CIF of relapse.

Finally, we should note that the Fine and Gray model for the subdistribution hazard $\lambda^{\mathrm{FG}}(t)$ and classical Cox models for the cause-specific hazards $\lambda_1(t)$ of relapse and $\lambda_2(t)$ of death are different models [42].

3 Regression Models

In this section we introduce recent regression models for the identiable quantities namely, cause specific hazard, cumulative incidence and conditional probability function. It should be pointed out that novel methodologies translate faster in medical journal. For exemple, the pseudo value approach introduced in 2003 is exemplified in a practical context in medical journal such as *Biology of Blood and Marrow Transplantation* [10] or [11].

3.1 Proportional hazards model

To relate the cause-specific hazard on the exposure covariate Z, the Cox proportional hazards model is often used while a similar model was proposed for the subdistribution hazard [13]. The Cox model expresses the cause-specific hazard as a multiplicative function of the baseline instantaneous hazard, $\lambda_{k0}(t) : \lambda_k(t) = \lambda_{k0}(t)\exp(\beta Z)$, where β is the covariate effect. The Fine and Gray model focuses on the hazard associated with the CIF and similarly expresses as $\alpha_k(t) = \alpha_{k0}(t)\exp(\gamma Z)$.

The Fine–Gray model, draw a lot of attention (from 2002) since its first use in HSCT. As a result several publications investigated the interpretation of the subdistribution hazard ratio [43]. The following references extand the Fine–Gray model or adapt standard methodologies to it [43, 44, 45, 46, 47, 36, 48]. Sun et al. [49] suggested a flexible additive multiplicative hazard model for modeling the subdistribution hazard.

In a recent paper Peng and Huang [50] propose a natural generalization of the Cox regression model, in which the regression coefficients have direct interpretations as temporal covariate effects on the survival function. Second-stage inferences with time-varying coefficients are developed accordingly. Simulations and a real example illustrate the practical utility of the proposed method.

3.2 Andersen–Klein model

A method based on pseudo-values has been proposed for direct regression modeling of the survival function [51, 52, 53, 19]. The pseudo value method is an estimating method. It enables the estimations of the following regression parameter, in a linear model for the CIF that was proposed by Fine [54]. The model for the CIF of type 1 is

$$g(F_1(t;Z)) = h(t) - Z\beta. \tag{1}$$

The parameter h(t) is the baseline failure probability, unspecified, invertible and stricly increasing in t. This general transformation model includes the Fine–Gray model taking $g(x) = log\{-log(1-x)\}$.

The Andersen–Klein model is an alternative estimation techniques for the model 1. Recently, in a series of papers, a method based on pseudo-values has been proposed for direct regression modeling of the survival function, the restricted mean and cumulative incidence function with right censored data.

$$g(F_1(t)) = F_{10}(t) + R(t)Z(t) \tag{2}$$

Note that this model encompasses time-dependent covariates through $Z(t)$ but requires that a grid or series of time points be specified. Usually 5 to 10 time points suffice to adequately model the CIF . The regression estimator of the parameter $R(t)$ is based on pseudovalues from the cumulative incidence function. Interestingly, the model (2), once the pseudo-values have been computed, can be fit using standard generalized estimating equation software. The use of these routines to obtain regression estimates for a study of bone marrow transplant patients is detailed in Klein et al. [53]. The model 2 is implented in the `pseudo` R–package.

Another appealling regression strategy is the Direct Binomial Regression [55] suggesting a new simple approach for estimation and assessment of covariate effects for the cumulative incidence curve in the competing risks model. They consider a semiparametric regression model where some effects may be time-varying and some may be constant over time. Their estimator can be implemented by standard software. Their simulation study shows that the estimator works well and has finite-sample properties comparable with the subdistribution approach. This methodology was exemplified to estimate the cumulative incidence of death in complete remission following a bone marrow transplantation. Interestingly, this regression model extends the Fine–Gray model, with time–dependent coefficients.

3.3 Time–dependent effects: The additive approach

A comprehensive description of additive model can be found in Martinussen and Scheike [15]. This class of model alternative is The Cox-Aalen additive-multiplicative intensity model that comprise a multiplicative part (Like a Cox model) and an additive part (like Aalen model). One interesting feature is that time-dependent effect and time-dependent covariate are easily handle indeed such properties violate the PH hazards assumptions.

Another important motivation for alternative modelling is pointed out in the recent work of Klein [56] the proportional hazard transition in multistate model can lead to inconsistencies. The additive models for either the hazard rates or the cumulative incidence functions are more *natural* and that these models properly partition the effect of a covariate on treatment failure into its component parts. These models are illustrated on data from a study of the efficacy of two preparative regimens for hematopoietic stem cell transplantation. Such findings must translate rapidly in HSCT.

Methods for fitting the Cox model with time–varying effects exist [57, 58, 59], but they all require some kind of smoothing thus depending on some smoothing parameter or sieve approximation. The obtained results may depend on the particular choice. The additive hazards regression model is an alternative (or supplement) to the Cox model. It was proposed by Aalen [60], and is is very flexible non parametric model. It results in plots that are informative regarding the effect of covariates on survival. The additive model of Aalen [60] specify the following relation betwen hazard and covariates :

$$\lambda_i(t) = \beta_0(t) + \beta_1(t)X_{i1}(t) + \ldots + \beta_p(t)X_{ip}(t)$$

An interesting submodel was suggested by Mckeague and Sasiene [61]

$$\lambda_i(t) = \exp\left(\beta(t)^T X_i(t) + \gamma^T Z_i(t)\right).$$

As pointed out by Klein [56], there is no guarantee that the estimated hazard is positive but this situation is very unlikely. The Martinussen–Scheike [62] model is a new additive-multiplicative hazard model which consists of two components. The first component contains additive covariate effects through an additive Aalen model while the second component contains multiplicative covariate effects through a Cox regression model. The Aalen model allows for time-varying covariate effects, while the Cox model allows only a common time-dependence through the baseline. This model is implemented in the `timereg` R–package.

3.4 Temporal process regression

This temporal process regression is a functional generalised linear model which specifies the mean of a response $Y(t)$ at time t conditionally on a vector of possibly timedependent covariates $Z(t)$, that is

$$E(Y(t)|\, Z(t)) = g^{-1}(\beta(t)' Z(t)), \tag{3}$$

where the link function g is monotone, differentiable and invertible. This is a very general model that encompass as particular case models such as logistic prevalence model. This model is implemented in the R–package tpr. A case study of this model can be found in the recent work of Allignol et al. [63] where this general regression framework was used to assess the effect of covariate on the conditional probability of a competing event [64].

4 Centre effect

A common question arising in multi-centre prospective clinical trials and in collaborative registry studies, is whether some heterogeneity in outcomes could be expected across centres, and, if such an heterogeneity exists, whether some statistical adjustment is required when estimating the prognostic effects of fixed covariates or not.

Recent developpments with an emphasis in HSCT are [65, 66, 45, 67]. Centre-effects are usually investigated with shared frailty models, and this presumes that this effect is constantly present during the follow-up, even when the follow-up is very long. More realistic are models with time-varying frailties. Therefore, the constant centre-specific frailty model was extended to allow time dependence of the frailties [66]. Notably, the center effect was adapted to the Fine–Gray model [45] introducing a random-effects model for the subdistribution hazard. This work was exemplified on data provided by the EBMT.

5 2–Sample Tests

In this section we introduce recent test statistics that are of great interest in the field of HSCT. Notably because, theses test have higher power to detect crossing *hazards*. In this section the terminology *hazard* will refer to CSH or SH. The major tests used in survival analysis are the log–rank test for comparing the equality of cause–specific hazards and the the Gray test for the comparison of subdistribution hazards. Crossing survival curves may be a consequence of crossing hazards and it is well known that for this situation many standard tests, such as the log-rank or Wilcoxon tests, will fail to pick up differences in survival curves [18]. Freidlin and Korn [35] formally compared the performance of log–rank and Gray's test. Small sample behaviour of variance estimator were investigated in Braun and Yuan [68]. Renyi type test was proposed as an alternative empowering users to detect differences between crossing hazards. It is a censored-data analogue of the Kolmogorov-Smirnov statistic and is based on the supremum of the absolute value of the entire path of the log-rank test statistic.

Bajorunaite and Klein [69] proposed a 2–sample tests for comparing cumulative incidence The test statistic is based on the maximum difference between two cumulative incidence functions and a second test based on the integrated weighted difference between the cumulative incidence functions for the event of interest in two samples (based on Pepe [70]'s test).

6 Sample size calculation

We have seen that numerous derivation of regression models are proposed for the analysis of effect of covariates. An essential step when planning a trial is the calculation of the sample size or the number of patients to recruit to detect a relevant effect with sufficient power. In HSCT, patients enrolled in a clinical trial may experience exclusive failure causes, which defines a competing risk setting. For instance, in hematology patients receiving a bone marrow transplantation may experience two exclusive events such as relapse and non-relapse death. Planning a trial when competing endpoints are acknowledged to exist thus requires appropriate methodology. Notably, when some (primary) endpoints rely on cumulative incidence inference, this must be accounted for when calculating a required number of events.

For both proportional cause–specific hazard and subdistribution hazards, sample size where derived in the presence of competing events [43, 71]. Both are based on Schoenfeld [72]'s formula and rely on similar key parameters, namely the hazard ratio that quantifies the treatment effect to be detected and the proportion of patients who are expected to fail from the cause of interest.

A sample size formula for the supremum log-rank test has also been recently presented in the classical survival framework [73]. It may be useful to anticipate possible departures from proportional hazards by using a test statistic less sensitive to this proportionality assumption. This is the case of Renyi-type tests also know as supremum log-rank tests in the classical survival framework [1416]. Recent work on sample size has shown that this test is nearly as efficient as the log-rank test when hazards are proportional, and can accommodate broader range of alternatives where the log-rank has no power to distinguish between groups. Additionally, Renyi-type test statistics have already been extended to the comparison of CIFs in the unpublished Ph.D. thesis of R. Bajorunaite. The Renyi-type tests are based on supremum integrated weighted difference of CIFs. We will refer to this test as the adapted Renyi-type test. More recently, Latouche and Porcher [47] suggested the use of supremum log–rank test and supremum Gray test .

7 Survival Analysis in R

A mandatory aspect for disseminating new statistical models is the availability of implementation. For exemple the `survival` package enables standard analysis for Cox model and Kaplan–Meier estimations [16].

To facilitate dissemination of R–package, we produced a *Task View* that enable user to easily install the whole packages related to survival analysis. To automatically install these views, the ctv package needs to be installed, e.g., via

```
>install.packages("ctv")
>library("ctv")
```

and then the views can be installed via install.views or update.views install.views(”Survival”) or update.views(”Survival”)

The Survival View is located at `http://cran.r-project.org/web/views/Survival.html`. This was done thanks to the collaboration of Arthur Allignol (Freiburg).

8 Conclusion

We have attempted to review recent developments in survival analysis and competing risks , with an emphasis on HSCT. The relevance of the use of such recent models are now established in the HSCT. In that respect the journal *Lifetime Data Analysis* has published a dedicated issue on Statistical analysis of HSCT Data [74, 75]. A question raised by this inventory, is to know whether or not recent developpments that bring new insights will reach applied statisticians/ clinicians. One solution would be to give statistical courses or educational session each year on a regular basis. The CLINT portal could be the core of this training/teaching infrastructure `http://clint.ebmt.org`.

Acknowledgments

The author would like to thank JP Klein for his comments on the early version of this work.

References

[1] T.M. Therneau and P.M. Grambsch. *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag Inc., New York, 2000.

[2] J.P. Klein and M. Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag Inc., London, 2003.

[3] J.D. Kalbfleish and R.L. Prentice. *The statistical analysis of failure time data*. Wiley, New York, 2002.

[4] D.W. Hosmer and S. Lemeshow. *Applied Survival Analysis: Regression Modeling of Time to Event Data*. John Wiley & Sons, Inc., New York, NY, USA, 1999. ISBN 0471154105.

[5] D.W. Hosmer, S. Lemeshow, and S. May. *Applied Survival Analysis: Regression Modeling of Time to Event Data, 2nd Edition*. John Wiley & Sons, Inc., New York, NY, USA, 2008. ISBN 978-0-471-75499-2.

[6] D. Collett. *Modelling survival data in medical research*. Chapman & Hall/CRC, Boca Raton, 2003.

[7] D.G. Kleinbaum and M. Klein. *Survival Analysis- A Self-Learning Text, Second Edition*. Springer-Verlag Inc., 2005. ISBN 978-0-387-23918-7.

[8] J.P. Klein, J.D. Rizzo, M.J. Zhang, and N. Keiding. Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part I: unadjusted analysis. *Bone Marrow Transplant*, 28(10):909–915, Nov 2001.

[9] J.P. Klein, J.D. Rizzo, M.J. Zhang, and N. Keiding. Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part 2: Regression modeling. *Bone Marrow Transplant*, 28(11):1001–1011, Dec 2001. Comparative Study.

[10] Brent R Logan, Mei-Jie Zhang, and John P Klein. Regression models for hazard rates versus cumulative incidence probabilities in hematopoietic cell transplantation data. *Biol Blood Marrow Transplant*, 12(1 Suppl 1):107–112, Jan 2006. Comparative Study.

[11] H.T. Kim. Cumulative incidence in competing risks data and competing risks regression analysis. *Clin Cancer Res*, 13(2 Pt 1):559–565, Jan 2007.

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

[13] J. P. Fine and R. J. Gray. A proportional hazards model for subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446):496–509, 1999.

[14] N. Balakrishnan and C.R. Rao, editors. *Advances in Survival Analysis*. Number 23 in Hanbook of statistics. Elsevier, Amsterdam, 2004.

[15] T. Martinussen and T.H. Scheike. *Dynamic Regression Models for Survival Data*. Springer-Verlag Inc., New York, 2006.

[16] S original by Terry Therneau and ported by Thomas Lumley. *survival: Survival analysis, including penalised likelihood.*, 2008. R package version 2.34.

[17] R. J. Gray. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics*, 116:1141–1154, 1988.

[18] J.P. Klein, B. Logan, M. Harhoff, and P.K. Andersen. Analyzing survival curves at a fixed point in time. *Statistics in Medicine*, 26:4505–4519, 2007.

[19] B.R. Logan, J.P. Klein, and M.J. Zhang. Comparing Treatments in the Presence of Crossing Survival Curves: An Application to Bone Marrow Transplantation. *Biometrics*, Jan 2008.

[20] P.K. Ruan and R.J. Gray. Sensitivity analysis of progression-free survival with dependent withdrawal. *Statistics in Medicine*, Sep 2007.

[21] J.P. Klein, N. Keiding, and E.A. Copelan. Plotting summary predictions in multistate survival models: probabilities of relapse and death in remission for bone marrow transplantation patients. *Statistics in Medicine*, 12:2315–2332, 1993.

[22] N. Keiding, J.P. Klein, and M.M. Horowitz. Multi-state models and outcome prediction in bone marrow transplantation. *Statistics in Medicine*, 20:1871–1885, 2001.

[23] J.P. Klein and C. Qian. Modeling multistate survival illustrated in bone marrow transplantation. Technical report, Medical College of Wisconsin, 1996.

[24] C Craddock, R M Szydlo, J P Klein, F Dazzi, E Olavarria, F van Rhee, C Pocock, K Cwynarski, J F Apperley, and J M Goldman. Estimating leukemia-free survival after allografting for chronic myeloid leukemia: a new method that takes into account patients who relapse and are restored to complete remission. *Blood*, 96(1):86–90, Jul 2000. Comparative Study.

[25] J.P. Klein, R.M. Szydlo, C. Craddock, and J.M. Goldman. Estimation of current leukaemia-free survival following donor lymphocyte infusion therapy for patients with leukaemia who relapse after allografting: application of a multistate model. *Statistics in Medicine*, 19(21):3005–3016, Nov 2000. Comparative Study.

[26] J.P. Klein, N. Keiding, Y. Shu, R.M. Szydlo, and J.M. Goldman. Summary curves for patients transplanted for chronic myeloid leukaemia salvaged by a donor lymphocyte infusion: the current leukaemia-free survival curve. *Br J Haematol*, 109(1):148–152, Apr 2000.

[27] P.K. Andersen, Ø. Borgan, R.D. Gill, and N. Keiding. *Statistical models based on counting processes.* Springer Series in Statistics. New York, NY: Springer., 1993.

[28] P. Hougaard. *Analysis of multivariate survival data*. Springer-Verlag Inc., 2000.

[29] P Hougaard. Multi-state models: A review. *Lifetime data analysis*, 5(3):239–264, 1999.

[30] PK Andersen and N Keiding. Multi-state models for event history analysis. *Statistical Methods in Medical Research*, 11(2):91–115, 2002.

[31] P.K. Andersen, S. Z. Abildstrøm, and S. Rosthøj. Competing risks as a multi-state model. *Statistical Methods in Medical Research*, 11(2):203–215, 2002.

[32] TH. Scheike and MJ Zhang. Direct modelling of regression effects for transition probabilities in multistate models. *Scandinavian Journal of Statistics*, 34(1):17–32, 2007.

[33] H. Putter, M. Fiocco, and R.B. Geskus. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine*, 26(11):2389–2430, 2007.

[34] L. Machado, J. Alvarez, Suarez C., and P.K. Andersen. Multi-state models for the analysis of time to event data. Technical report, Department of Biostatistics, University of Copenhagen., 2007.

[35] B. Freidlin and E.L. Korn. Testing treatment effects in the presence of competing risks. *Statistics in Medicine*, 24(11):1703–1712, 2005.

[36] J. Beyersmann and M. Schumacher. Misspecified regression model for the subdistribution hazard of a competing risk. *Statistics in Medicine*, 26(7):1649–1651, 2007. Comment.

[37] G. Bakoyannis and Touloumi G. A practical guide on modeling competing risk data. Technical report, On behalf of CASCADE Collaboration., 2008.

[38] E. J. Shpall, R. Quinones, R. Giller, C. Zeng, A. E. Baron, R. B. Jones, S. I. Bearman, Y. Nieto, B. Freed, N. Madinger, C. J. Hogan, V. Slat-Vasquez, P. Russell, B. Blunk, D. Schissel, E. Hild, J. Malcolm, W. Ward, and I. K. McNiece. Transplantation of ex vivo expanded cord blood. *Biol. Blood Marrow Transplant.*, 8:368–376, 2002.

[39] E. Meyer, J. Beyersmann, H. Bertz, S. Wenzler-Rttele, R. Babikir, M. Schumacher, F.D. Daschner, H. Rden, and M. Dettenkofer. Risk factor analysis of blood stream infection and pneumonia in neutropenic patients after peripheral blood stem-cell transplantation. *Bone Marrow Transplant.*, 39:173–178, 2007.

[40] Odd O. Aalen and Håkon K. Gjessing. Understanding the shape of the hazard rate: A process point of view. (With comments and a rejoinder). *Stat. Sci.*, 16(1):1–22, 2001.

[41] J.D. Kalbfleisch and R.L. Prentice. *The statistical analysis of failure time data. 2nd ed.* Wiley Series in Probability and Statistics. Chichester: Wiley., 2002.

[42] A. Latouche, V. Boisson, S. Chevret, and R. Porcher. Misspecified regression model for the subdistribution hazard of a competing risk. *Statistics in Medicine*, 26(5):965–974, Feb 2007.

[43] A. Latouche, R. Porcher, and S. Chevret. Sample size formula for proportional hazards modelling of competing risks. *Statistics in Medicine*, 23(21):3263–3274, 2004.

[44] A. Latouche, R. Porcher, and S. Chevret. A note on including time-dependent covariate in regression model for competing risks data. *Biometrical Journal*, 47:807–814, 2005.

[45] S. Katsahian, M. Resche-Rigon, S. Chevret, and R. Porcher. Analysing multicentre competing risks data with a mixed proportional hazards model for the subdistribution. *Statistics in Medicine*, 25:4267–4278, Dec 2006.

[46] A. Latouche, J. Beyersmann, and J.P. Fine. Comments on 'Analysing and interpreting competing risk data'. *Statistics in Medicine*, 26:3676–3679, Aug 2007.

[47] A. Latouche and R. Porcher. Sample size calculations in the presence of competing risks. *Statistics in Medicine*, 26:5370–5380, Dec 2007.

[48] J. Beyersmann, M. Dettenkofer, H. Bertz, and M. Schumacher. A competing risks analysis of bloodstream infection after stem-cell transplantation using subdistribution hazards and cause-specific hazards. *Statistics in Medicine*, 2007.

[49] L. Sun, J. Liu, J. Sun, and Zhang M.J. Modeling the subdistribution of a competing risks. *Statistica Sinica*, pages 1367–1385, 2006.

[50] Limin Peng and Yijian Huang. Survival analysis with temporal covariate effects. *Biometrika*, 94 (3):719–733, 2007.

[51] P. K. Andersen, J. P. Klein, and S. Rosthøj. Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika*, 90(1):15–27, 2003.

[52] P. K. Andersen and J.P. Klein. Regression analysis for multistate models based on a pseudo-value approach, with applications to bone marrow transplantation studies. *Scandinavian Journal of Statistics*, 34(1):3–16, 2007.

[53] J.P. Klein, M. Gerster, P.K. Andersen, S. Tarima, and M.P. Perme. SAS and R functions to compute pseudo-values for censored data regression. *Comput Methods Programs Biomed*, 89: 289–300, Mar 2008.

[54] J. P. Fine. Regression modelling of competing crude failure probabilities. *Biostatistics*, 2:85–97, 2001.

[55] T.H. Scheike, Mei-Jie Zhang, and T.A. Gerds. Predicting cumulative incidence probability by direct binomial regression. *Biometrika*, 95(1):205–220, 2008.

[56] J.P. Klein. Modelling competing risks in cancer studies. *Statistics in Medicine*, 25(6):1015–1034, 2006.

[57] T Martinussen, TH Scheike, and IM Skovgaard. Efficient estimation of fixed and time-varying covariate effects in multiplicative intensity models. *Scandinavian Journal of Statistics*, 29(1): 57–74, 2002.

[58] Z. Cai and Y. Sun. Local linear estimation for time-dependent coefficients in cox's regression models. *Scandinavian Journal of Statistics*, 30(1):93–111, 2003.

[59] L Tian, D Zucker, and LJ Wei. On the cox model with time-varying regression coefficients. *Journal of the American Statistical Association*, 100(469):172–183, 2005.

[60] O.O. Aalen. A linear regression model for the analysis of life times. *Statistics in Medicine*, 8: 907–925, 1989.

[61] I.W. Mckeague and P.D. Sasiene. A partly parametric additive risk model. *Biometrika*, 81(3): 501–514, 1994.

[62] T. Martinussen and T.H. Scheike. A flexible additive multiplicative hazard model. *Biometrika*, 89(2):283–298, 2002.

[63] A. Allignol, A. Latouche, J. Yan, and J.P. Fine. A regression model for the conditional probability of a competing event: Application to monoclonal gam- mopathy of unknown significance. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, In press, 2010.

[64] A. Allignol. *Cprob: Conditional probability function of a competing event*, 2010. URL `http://CRAN.R-project.org/package=Cprob`. R package version 1.1.

[65] T. Yamaguchi, Y. Ohashi, and Y. Matsuyama. Proportional hazards models with random effects to examine centre effects in multicentre cancer clinical trials. *Stat Methods Med Res*, 11:221–236, 2002.

[66] C. Claire Wintrebert, H. Hein Putter, A. Zwinderman, and J.C. Houwelingen. Centre-effect on survival after bone marrow transplantation: Application of time-dependent frailty models. *Biometrical Journal*, 46:512–525, 2004.

[67] D.V. Glidden and E. Vittinghoff. Modelling clustered survival data from multicentre clinical trials. *Statistics in Medicine*, 23:369–388, 2004.

[68] T.M. Braun and Z. Yuan. Comparing the small sample performance of several variance estimators under competing risks. *Statistics in Medicine*, 26:1170–1180, 2007.

[69] R. Bajorunaite and J.P. Klein. Two-sample tests of the equality of two cumulative incidence functions. *Comput. Stat. Data Anal.*, 51(9):4269–4281, 2007.

[70] M.S. Pepe. Inference for events with dependent risks in multiple endpoint studies. *Journal of the American Statistical Association*, 86(415):770–778, 1991.

[71] G. Schulgen, M. Olschewski, V. Krane, C. Wanner, G. Ruf, and M. Schumacher. Sample sizes for clinical trials with time-to-event endpoints and competing risks. *Contemporary Clinical Trials*, 26(3):386–396, 2005.

[72] D. A. Schoenfeld. Sample-size formula for the proportionnal-hazards regression model. *Biometrics*, 39:499–503, 1983.

[73] K.H. Eng and M. R. Kosorok. A sample size formula for the supremum log-rank statistic. *Biometrics*, 61(1):86–91, 2005.

[74] M. Eapen and V. Rocha. Principles and analysis of hematopoietic stem cell transplantation outcomes: the physician's perspective. *Lifetime Data Anal*, 14:379–388, 2008.

[75] T.H. Scheike and M.J. Zhang. Flexible competing risks regression modeling and goodness-of-fit. *Lifetime Data Anal*, 14:464–483, 2008.