ABSOLUTE RISK REGRESSION FOR COMPETING RISKS: INTERPRETATION, LINK FUNCTIONS AND PREDICTION

THOMAS A. GERDS[1], **THOMAS H. SCHEIKE**[1], and **PER K. ANDERSEN**[1]

[1] Department of Biostatistics, University of Copenhagen, Denmark

Abstract

In survival analysis with competing risks the transformation model allows different functions between the outcome and explanatory variables. However, the model's prediction accuracy and the interpretation of parameters may be sensitive to the choice of link function. We review the practical implications of different link functions for regression of the absolute risk (or cumulative incidence) of an event. Specifically we consider models in which the regression coefficients β have the following interpretation: The probability of dying from cause D during the next t years changes with a factor $\exp(\beta)$ for a one unit change of the corresponding predictor variable, given fixed values for the other predictor variables. The models have a direct interpretation for the predictive ability of the risk factors. We propose some tools to justify the models in comparison with traditional approaches which combine a series of cause-specific Cox regression models, or use the Fine-Gray model. The methods are illustrated using bone marrow transplant data.

Keywords

1. Introduction

Competing risks models are nowadays in routine use for the analysis of clinical trials and epidemiological studies. A boom in biomarker research and the aim to predict the future disease course of patients have increased the demand for statistical methods that quantify the predictive ability of genotype, phenotype, treatment and environmental factors. For example, a patient diagnosed with diabetes may be interested in the risk of cardiovascular disease related death. In a broader context, it is of interest to quantify how multiple risk factors change the predicted risk of death caused by cardiovascular disease. In this article we review and compare the practical properties of different regression models for competing risks, specifically for predicting the individual risks of cancer patients (Figure 1).

In applications the choice of a prediction model relies on a number of considerations, including:

- Size of prediction error

- Fit of model

- Interpretation of model parameters

- Mathematical coherence of models

The estimation of absolute risk based on estimates of the cause-specific hazard functions was discussed for example in [1, 2]. Unfortunately, hazard ratios as obtained by cause-specific Cox regression analyses do not directly quantify the ability of the single markers to predict the unconditional absolute risk of an event of interest. The Fine-Gray model [3] can be used to test if the cumulative incidence depends on a risk factor, but the absolute values of the resulting regression coefficients are difficult to interpret [3, 4]. We specifically discuss multiple regression models that directly quantify the expected change of the predicted absolute risk of an event (cumulative incidence) for a one unit change of one predictor's value given fixed values for the other predictor variables. Such models include the ratio between cumulative incidences in two groups as the special case with exactly one binary predictor variable. Recently, Zhang et al. [5] investigated the difference, the odds ratio and also the ratio between two cumulative incidence functions. Another useful model is the logistic risk regression model which is an extension of the odds ratio to multiple regression in competing risks. However, relative absolute risks are easier to understand [6].

These models are not new and the mathematical properties are well-studied in the context of the linear transformation model [7]. The reasons why alternatives to cause-specific Cox regression and Fine-Gray regression are not used in practice or discussed in monographs on competing risks models are not clear. From a purely mathematical viewpoint the absolute risk regression models have the problem that their formulation does not guarantee that the predicted probability of an event is between 0% and 100%. However, we exemplify in this article that the predictions can be as good as those from the Fine-Gray model and from combined cause-specific Cox regression analyses. The logistic link yields mathematically coherent models, but in the presence of competing risks the interpretation of the resulting odds ratios maybe more cumbersome than usual, because one minus the absolute risk of an event is the sum of the survival chance and the risk of competing events.

We present a worked example where we consider data from patients who received a bone marrow transplant. We compare the predictive ability of different models using scatterplots at single time-points, and via cross-validated time-dependent mean squared error of prediction adapted to censored data and competing risks [8-10].

For the convenience of the reader we provide the library *riskRegression*, a collection of user-friendly R-functions to fit the different models considered in this article. Sample code is provided in an online appendix.

2. Competing risks regression

In a competing risks framework, let T be the exit time from the initial state to one of the absorbing states (events), and D indicate the type of event [11]. We are interested in regression models for the cumulative incidence function F_1 whose values are the time-dependent absolute probabilities of occurrence of event 1:

$$F_1(t|X) = P(T \le t, D=1|X) = \int_0^t S(s-|X)\,\alpha_1(s|X) \quad \mathrm{d}s. \quad (1)$$

Here $X = (X_1, \ldots, X_K)$ is a vector of predictor variables measured at baseline for subject i, a_1 is the cause-specific hazard function for event 1 (Figure 1), and S denotes the event-free survival function. More precisely, $S(t|X)$ is the time-dependent probability that a patient with risk factors X stays in the initial state until time t.

A desirable conclusion sentence from a regression analysis of F_1 would be: "The probability of dying from cardiovascular disease during the next t years is $\exp(\beta)$ times as high for a patient with diabetes than for a patient without diabetes, given fixed values for the other predictor variables." The regression parameters $\beta = (\beta_1, \ldots, \beta_K)$ in the following absolute risk regression model (ARR) have the desired interpretation:

$$F_1^{(arr)}(t|X) = F_{1,0}(t)\ exp\left(\beta_1 X_1 + \cdots \beta_K X_K\right). \quad (2)$$

Here $F_{1,0}$ is an unspecified function of time which represents the cumulative incidence for subjects with $X = 0$. The logarithmic link function yields that predictor variable changes are translated into ratios of the corresponding cumulative incidences:

$$\frac{F_1^{(arr)}(t|X_1=x_1,\ldots,X_k=x_k+1,\ldots,X_K=x_K)}{F_1^{(arr)}(t|X_1=x_1,\ldots,X_k=x_k,\ldots,X_K=x_K)} = exp\left(\beta_k\right). \quad (3)$$

Consistent estimators of the model parameters are defined as solutions to generalized estimation equations [12] suitably adapted for censored data and competing risks. A corresponding semi-parametric theory is available for the transformation model:

$$g\left\{F_1(t|X)\right\} = \beta_0(t) + \beta_1 X_1 + \cdots \beta_K X_K \quad (4)$$

where g is a known differentiable function. The transformation model includes the model (2) as the special case where $g(p) = \log(p)$ and $\beta_0(t) = \log F_{1,0}(t)$. The logistic-link model corresponds to $g(p) = \log(p/(1-p))$ and the Fine-Gray model to the complementary log-log link: $g(p) = \log(-\log(p))$. The properties of the transformation model have been studied by many [3,13-17]. Specifically, to deal with right censored data, estimation techniques are available based on inverse of the probability of censoring weights (IPCW) [3, 15, 17] or jackknife pseudo-values [18].

3. Illustration: Bone marrow transplant study

For the purpose of illustrating the methods we consider the data from 1715 leukemia patients who received a bone marrow transplant (BMT) [19]. The endpoints are what comes first: relapse of the disease or death in remission (Figure 1). A total of 847 patients survived the follow-up period in remission, follow-up ended in relapse for 311 patients and in death in remission for 557 patients. According to the (reverse) Kaplan-Meier for the censoring times the median follow-up time was 37.2 months (inter quartile range: [23.7,52.7]), see Figure 2.

The following predictor variables were used to predict the absolute risks of relapse and death: disease type (levels: ALL, AML, CML), time waiting for transplant in months since

diagnosis (median: 10.95 months; IQR: [6,23]; Range: [0.4;200]), patient gender (male: 985; female:727), dichotomized Karnofsky index (positive: 1382; negative: 333), disease stage (early: 1026; intermediate: 410; advanced: 279), and type of donor (sibling:1224; matched/ unrelated:383; mismatched/unrelated:108).

3.1. Inverse of the probability of censoring weights

For the purpose of using the IPCW methods for fitting direct regression models as described in [17] we investigated the association between censoring times and predictor variables. Based on a Cox regression model for the censoring times, Table 1 shows significant effects of several predictor variables. Positive Karnofsky index, disease type AML (compared to disease type ALL), and intermediate disease stage (compared to early disease stage) have a negative effect on the censoring hazard and indicate longer follow-up, and a matched/ unrelated donor versus a sibling donor has a positive effect on the censoring hazard and indicates a shorter follow-up period given fixed values for the other predictors.

This result leads to the well-known dilemma of the IPCW technique: Ignoring the covariates and using weights based on the marginal Kaplan-Meier estimate of the censoring distribution may introduce a bias if there is a true relation between the predictors and the censoring distribution, and will not yield efficient estimates because the predictor values of the patients that were lost to follow-up (eventfree) will not enter into the statistic. On the other hand, if a parametric model, like the Cox regression model presented in Table 1, is biased then the results of IPCW may also be biased. Note that a purely nonparametric approach is often not feasible due to the curse of dimensionality. In the BMT data considered here, if the continuous predictor "transplant waiting time" is ignored, then a stratified Kaplan-Meier estimate could be used for the IPCW technique. However, all possible combinations of the values of the variables disease stage and type, Karnofsky index and donor matching generate 54 different classes of which 12 classes would include less than 4 patients. Hence, we do not consider nonparametric weights and in all what follows below work with IPC weights derived from the Cox regression model (Table 1).

3.2. Regression models

The following regression analyses were performed to predict the event probabilities: 1. Absolute risk regression, 2. Log-odds regression, 3. Fine-Gray regression, 4. Combination of cause-specific Cox regression. For each of the four approaches we separately performed one regression analysis for each of the two endpoints, *relapse* and *death in remission*. For each model we included the same predictor variables in additive form and did not consider interactions. The models 1-3 were implemented in the statistical software R [20] using IPCW based on weights derived from the Cox regression model presented in section 3.1 and Table 1. For details on the implementation see [21]. Note that the combination of the cause-specific Cox regression models can be estimated using partial likelihood, and does not require a model for the censoring distribution. Instead one has to specify models for the cause-specific hazard functions of all competing risks.

Table 2 shows the results of the absolute risks regression models. A sample conclusion sentence derived from the table is the following:

"A positive Karnofsky index significantly increased the absolute risk of relapse, by a factor 1.3 (95%-CI=[1.01;1.65]), and significantly decreased the risk of dying in remission, by a factor 0.81 (95%-CI=[0.7;0.94]).".

Table 3 shows results obtained with different multiple regression models for the effects of the factor donor on the two competing risks.

The parameters $exp\left(\hat{\beta}\right)$ obtained with the logistic-link regression models can be interpreted as ratios between the odds of experiencing the event of interest. However, the complementary probability of "not experiencing the event" includes both the chance of no event and the risk of the competing event. As usual the absolute deviations of the odds ratios from the reference value 1 are systematically higher compared to absolute relative risks.

The values of the estimated parameters obtained with the Fine-Gray regression model are difficult to interpret within the context of the proportional hazard model for the sub-distribution hazard function [3, 4].

However, the first three models agree with respect to statistical significance of all four parameters shown in Table 3. The effect of matched/unrelated donor vs. sibling quantified by an estimated cause-specific hazard ratio of 1.05 (95% confidence interval: [0.78;1.42]) is not statistically significant ($p > 0.05$) in the cause-specific Cox regression model for relapse. This observation has been documented elsewhere [18, 22].

4. Model checking

The models discussed in the previous section all assumed that the regression effects do not depend on time.

To explore this assumption we consider an extension of the models that allow time-interaction [23]. In the absolute risk context a model that captures a potential time-interaction is

$$F_1^{tarr}\left(t, X\right) = F_{1,0}\left(t\right) \; exp\left\{\gamma_1\left(t\right) X_1 + \cdots + \gamma_K\left(t\right) X_K\right\}$$

where the regression coefficients γ_k are functions of time. Fitting this type of model for the absolute risk, logistic, and Fine-Gray models reveals that all models have severe problems with the disease stage variable. This is illustrated in Figure 3 which shows the time-varying effect of intermediate vs early disease stage for the absolute risk model. All other covariates can approximately be described as having constant effects (not shown). Figure 3 also shows the estimate of the effect of matched/unrelated donors compared to sibling donors for the absolute risk regression model, that (for all link functions) indicate some slight time-interaction. A formal test can be constructed as described in detail in [24]. The time-varying coefficient model is used here for the purpose of model checking, and it is important to note that estimates may not be be valid subdistributions. Another way of examining whether the effect (on the proper scale) is time-constant is to use pseudo-observations as discussed in [25].

To fully comply with the assumption that the effects do not change with time, one possibility is to restrict attention to a different time-interval. The figures can be translated into p-values using resampling techniques [23] but the key is really to consider the graphical display of the effects.

5. Prediction accuracy

To justify the absolute risk regression model, in particular the link function, a good starting point is to compare the predictions and prediction accuracy to that obtained with established models. A simple first check is to compare the predictions of event probabilities at a given time-point.

Using the BMT data and models introduced in section 3, the scatterplots comparing the predicted risks of relapse, respectively of death in remission, do not show great differences of individual predictions (Figure 4). The time-fixed coefficient absolute risk regression model predicted on average (over the 1715 patients in the BMT data) a 0.48 % (min: −10.2%; max: 5.5%) higher probability of relapse after 3 years than the cause-specific Cox regression model and a 0.3% (min: −2.6%; max: 3.1%) higher probability of relapse after 3 years than the Fine-Gray regression model. The differences were slightly larger for predicting death in remission than for predicting relapse. Here the time-fixed coefficient absolute risk regression model predicted on average a 0.04 % (min: −15.9%; max: 12.4%) lower probability than the cause-specific Cox regression model and a 0.04% (min: −16.5%; max: 7.9%) lower probability than the Fine-Gray regression model.

The next question is whether these deviations in predicted probabilities differentiate the prediction accuracy of the alternative models. The time-dependent Brier score can be estimated separately for the different events using a weighted average of individual residuals [26, 10, 27]:

$$BS\left(t, F_1\right) = \frac{1}{N} \sum_{i=1}^{N} W\left(t, X\right) \left\{ N_{1i}\left(t\right) - F_1\left(t|X\right) \right\}^2.$$

Here, inverse of the probability of censoring weights $W\left(t, X\right)$ are based on an estimate of the conditional censoring distribution given the predictor variables [10]. Again, these weights are derived from the Cox regression model shown in Table 1.

Figure 5 compares the estimated Brier scores for the absolute risk regression model with time-fixed coefficients to Fine-Gray regression and cause-specific Cox regression. The analysis shows that the three models have quite similar prediction performance. For example, the Brier scores for predicting relapse during the first three years are estimated as 12.73%, 12.64% and 12.72% for the absolute risk regression, the Fine-Gray regression and the cause-specific Cox regression, respectively.

6. Discussion

Prediction of absolute risk is an important activity in many disciplines, including medicine. In this paper we have studied a number of different regression models for prediction of relapse and of death in remission in a bone marrow transplantation setting (Tables 2 and 3). In this data set we reviewed the choice of link function and observed that, in terms of individual predictions (Figure 4) and prediction error (Figure 5), no substantial differences were identified among the models. This was in spite of the fact that for all direct models, deviations from the assumption of time-constant effects (on the appropriate scale) were seen (Figure 3).

When it comes to parameter interpretation, some of the fitted models are more attractive than others. The two standard competing risks regression models are Cox models for cause-specific hazards and the Fine-Gray regression model. For the former, $\exp(\beta)$ parameters have standard rate ratio interpretations as known from epidemiology. However, in a competing risks model rate ratios do not directly translate to relationships between risks (cumulative incidences) [28] as also illustrated for the effect of donor in Table 3. For the Fine-Gray model, $\exp(\beta)$ parameters are "subdistribution hazard ratios" and since a subdistribution hazard (being the rate of event among those who are *either* still alive *or* have already died from a competing cause) has a quite indirect interpretation, so have the regression coefficients from that model. However, the Fine-Gray model does establish a useful direct link between covariates and cumulative incidence.

This link is also evident for the other two models fitted to the bone marrow transplantation data in Section 3. Thus, for the logit link model, $\exp(\beta)$ parameters are ratios between odds of the form

$$F_1\left(t|X\right)/\left(1 - F_1\left(t|X\right)\right),$$

i.e. the cumulative incidence divided by its complement, the probability of either having survived or having died from a competing cause by time *t*. Such odds parameters may suffer from some of the same drawbacks as the subdistribution hazard and below, we discuss a possible model for an alternative odds parameter. For the log link model we argue that the parameter interpretation is simple and useful: $\exp(\beta)$ parameters are ratios between cumulative incidences.

So, why is the log link model not the standard direct regression model for a cumulative incidence? One reason may have to do with the mathematical properties of the model. Probabilities from the model may exceed 1. How large a practical problem this may be is hard to say in general and in our view this is most likely to be a problem if unwarranted extrapolations beyond the range of observed covariates are aimed at. On the more technical side, some care must be exercised when fitting the model due to numerical instabilities for small values of time *t* arising from the fact that the log cumulative incidence is undefined for *t* = 0. This is also a problem for the logit link model but not for the Fine-Gray model.

However, a problem possessed by all direct regression models for cumulative incidences is that the sum of all predicted cumulative incidences, that is, $\hat{F}_1(t|X) + \hat{F}_2(t|X)$ for two competing causes may exceed 1. If focus is on a single cause then one might argue that this is a minor problem but for a thorough competing risks analysis, all causes should be studied and the problem does become relevant. This problem does not occur for cause-specific hazard (Cox or other) models for which predicted probabilities add up to 1:

$$\hat{F}_1(t|X) + \hat{F}_2(t|X) + \hat{S}(t|X) = 1.$$

The following alternative logit model also has this desirable property. Consider the case with two competing causes and assume that

$$log\frac{F_j(t|X)}{S(t|X)} = A_j(t) + X^{\mathsf{T}}\beta^j, \; j=1,2.$$

This is a continuous time multinomial logistic model (see e.g., Hosmer and Lemeshow [29, Sect. 8.1] for a presentation of the standard multinomial logistic model) and the three probabilities

$$F_1(t|X) = \frac{exp\left(A_1(t) + X^{\mathsf{T}}\beta^1\right)}{1 + exp\left(A_1(t) + X^{\mathsf{T}}\beta^1\right) + exp\left(A_2(t) + X^{\mathsf{T}}\beta^2\right)},$$

the corresponding $F_2(t \mid X)$ and

$$S(t|X) = \frac{1}{1 + exp\left(A_1(t) + X^{\mathsf{T}}\beta^1\right) + exp\left(A_2(t) + X^{\mathsf{T}}\beta^2\right)},$$

do add up to 1. The $exp(\beta^j)$ parameters from this model have a simpler interpretation than those from the logit model fitted in section 3. Thus, $exp(\beta_j)$ parameters are ratios between "odds" of the form $F_j(t \mid X)/S(t \mid X)$, that is, the risk of a cause j failure in relation to the probability of no failure. A drawback, however, is that (like cumulative incidences predicted from cause-specific hazards) the cause 1 cumulative incidence predicted from this logit model depends on both the β^1 parameters and the corresponding parameters, β^2, for the competing cause.

How to balance the criteria for choosing among competing prediction models in a given situation may not be obvious. We have discussed some of the considerations which we have found relevant and this has led us to the conclusion that, in spite of its mathematical inconveniences, the log link cumulative incidence regression model (2) due to its desirable parameter interpretation may be a serious competitor to more standard models.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Benichou J, Gail MH. Estimates of absolute cause-specific risk in cohort studies. Biometrics. 1990; 46(3):813–826. [PubMed: 2242416]

2. Korn EL, Dorey FJ. Applications of crude incidence curves. Statistics in medicine. 1992; 11(6): 813–829. [PubMed: 1594819]

3. Jason P. Fine and Robert J. Gray. A proportional hazards model for the subdistribution of a competing risk. J. Am. Stat. Assoc. 1999; 94(446):496–509.

4. Andersen, PK.; Keiding, N. Interpretability and importance of functionals in competing risks and multistate models. Technical Report 10/6. Department of Biostatistics, University of Copenhagen; 2010.

5. Zhang, Mei-Jie; Fine, Jason. Summarizing di erences in cumulative incidence functions. Stat Med. 2008; 27(24):4939–49. [PubMed: 18563792]

6. Marschner IC, Gillett AC. Relative risk regression: reliable and flexible methods for log-binomial models. Biostatistics. 2012; 13(1):179–192. [PubMed: 21914729]

7. Fine JP. Analysing competing risks data with transformation models. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 1999; 61(4):817–830.

8. Gerds TA, Schumacher M. Efron-type measures of prediction error for survival analysis. Biometrics. 2007; 63(4):1283–1287. [PubMed: 17651459]

9. Rosthøj, S.; Keiding, N. Explained variation and predictive accuracy with an extension to the competing risks model. Technical Report 03/14. Department of Biostatistics, University of Copenhagen; 2003.

10. Binder H, Allignol A, Schumacher M, Beyersmann J. Boosting for high-dimensional time-to-event data with competing risks. Bioinformatics. 2009; 25(7):890–896. [PubMed: 19244389]

11. Andersen PK, Abildstrom SZ, Rosthøj S. Competing risks as a multi-state model. Statistical methods in medical research. 2002; 11:203–205. [PubMed: 12040697]

12. Liang, Kung-Yee; Zeger, Scott L. Longitudinal data analysis using generalized linear models. Biometrika. 1986; 73:13–22.

13. Fine JP, Ying Z, Wei LG. On the linear transformation model for censored data. Biometrika. 85(4): 980.

14. Fine JP. Analysing competing risks data with transformation models. Journal Royal Statistical Society: Series B. 1999; 61(4):817–830.

15. Fine, Jason P. Regression modeling of competing crude failure probabilities. Biostatistics. 2001; 2(1):85–97. [PubMed: 12933558]

16. Fine JP, Yan J, Kosorok MR. Temporal process regression. Biometrika. 91(3):683.

17. Scheike, Thomas H.; Zhang, Mei-Jie; Gerds, Thomas A. Predicting cumulative incidence probability by direct binomial regression. Biometrika. 2008; 95:205–220.

18. Klein, John P.; Andersen, Per Kragh. Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. Biometrics. 2005; 61(1):223–229. [PubMed: 15737097]

19. Klein JP, Gale RP, Ash RC, Bach FH, Bradley BA, Casper JT, Flomenberg N, Gajewski JL, Gluckman E, Henslee-Downey PJ, Hows JM, Jacobsen N, Kolb HJ, Lowenberg B, Masaoka T, Rowlings PA, Sondel PM, van Bekkum DW, van Rood JJ, Vowels MR, Zhang MJ, Horowitz MM, Szydlo R, Goldman JM. Results of allogeneic bone marrow transplants for leukemia using donors other than hla-identical siblings. Journal of Clinical Oncology. 1997; 15(5):1767. [PubMed: 9164184]

20. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2009. ISBN 3-900051-07-0

21. Scheike TH, Zhang MJ. Analyzing competing risk data using the R timereg package. Journal of Statistical Software. 2011; 38(2):1–15.

22. Andersen PK, Klein JP. Regression analysis for multistate models based on a pseudo-value approach, with applications to bone marrow transplantation studies. Scandinavian Journal of Statistics. 2007; 34:3–16.

23. Scheike TH, Zhang M-J. Flexible competing risks regression modelling and goodness-of-fit. Lifetime Data Analysis. 2008; 14:464–483. [PubMed: 18752067]

24. Scheike T, Zhang M. Flexible competing risks regression modeling and goodness-of-fit. Lifetime Data Analysis. 2008; 14(4):464–83. [PubMed: 18752067]

25. Andersen PK, Perme MP. Pseudo-observations in survival analysis. Statistical Methods in Medical Research. 2010; 19(1):71–99. [PubMed: 19654170]

26. Gerds TA, Schumacher M. Consistent estimation of the expected Brier score in general survival models with right-censored event times. Biometrical Journal. 2006; 48:1029–1040. [PubMed: 17240660]

27. Schoop R, Beyersmann J, Schumacher M, Binder H. Quantifying the predictive accuracy of time-to-event models in the presence of competing risks. Biometrical Journal. 2011; 53:88–112. [PubMed: 21259311]

28. Andersen, PK.; Geskus, R.; Putter, H. Competing risks in epidemiology: Possibilities and pitfalls. Technical Report 2. Department of Biostatistics, University of Copenhagen; 2011.

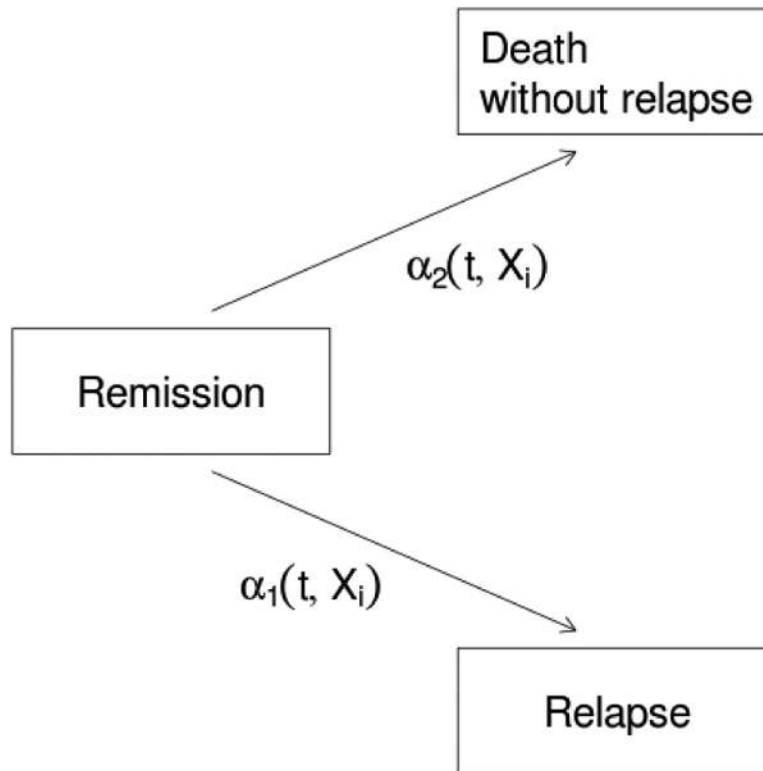29. Hosmer, DW.; Lemeshow, S. Applied logistic regression. Vol. 354. Wiley-Interscience; 2000.

Figure 1.
A competing risks model describes the time course of subjects that share a common initial state at the time origin (Remission). The time course is terminated when either of the competing events (Event 1: relapse or Event 2: death without relapse) has occurred. The cause-specific hazard functions α_1 and α_2 describe the instantaneous rates of the two events at time t.
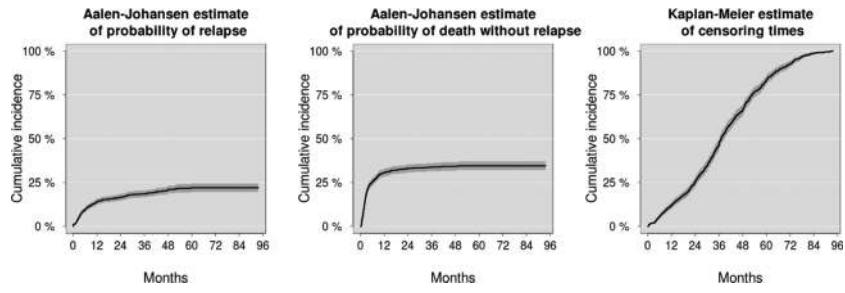
Figure 2.
Cumulative incidences of events and estimate of follow-up distribution for the BMT study
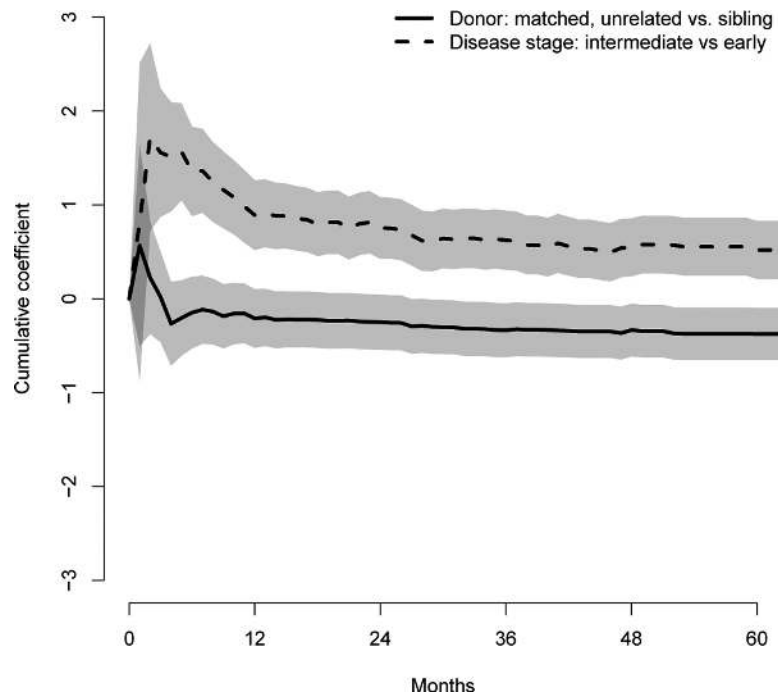
Figure 3.
Time-dependent effects in the absolute risk regression model for relapse. The non-parametric estimates are shown with 95% pointwise confidence intervals.
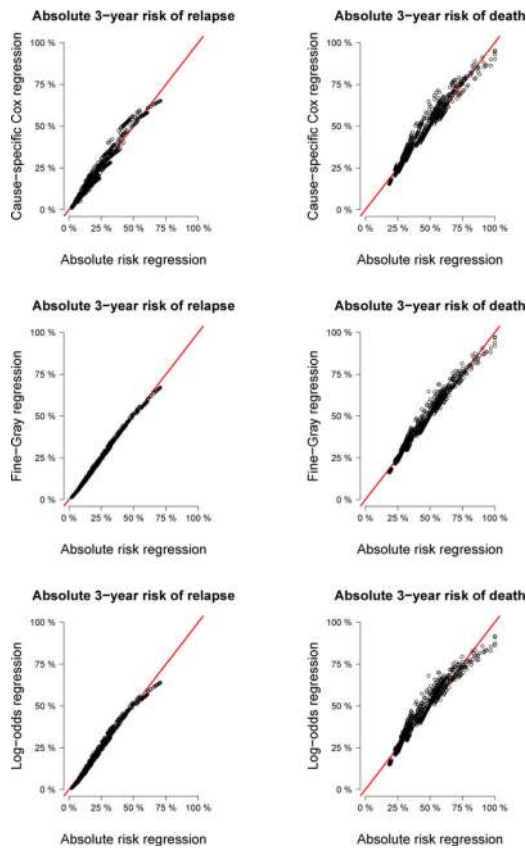
Figure 4.
Predicted cumulative probabilities for the patient status 3 years after the transplant. Compared are the predictions for relapse and death in remission based on absolute risk regression (x-axes) and Fine-Gray regression (y-axes top panels), cause-specific Cox regression (y-axes middle panels), log-odds regression (y-axes bottom panels). Each of the black dots represents the predictions at the predictor values of one patient in the BMT data set.
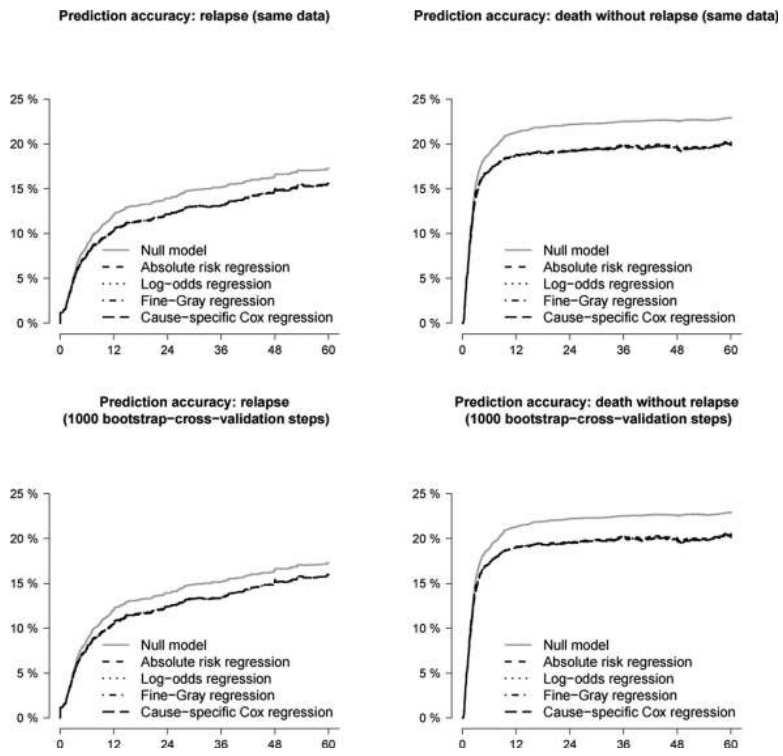
Figure 5.
Prediction error (Brier score) estimated for the predictions for relapse and death in remission based on the same data used for fitting the models (top panels) and based on 1000 steps of bootstrap cross-validation (bottom panels).

Table 1

Results from Cox regression of the observed censoring times.

	Hazard ratio	95% CI	P-value
disease:ALL	–	–	–
disease:AML	0.79	[0.63; 0.97]	0.026
disease:CML	0.95	[0.78; 1.15]	0.58
Karnofsky	0.74	[0.60; 0.92]	0.0076
donor:sibling	–	–	–
matched/unrelated	1.75	[1.44; 2.13]	< 0.0001
mismatched/unrelated	1.02	[0.69; 1.50]	0.94
stage:early	–	–	–
stage:intermediate	0.78	[0.64; 0.95]	0.011
stage:advanced	1.29	[0.96; 1.75]	0.094
transplant waiting time	1.00	[0.99; 1.00]	0.40

Table 2

Results from absolute risk regression.

Factor	Relapse			Death in remission		
	Relative absolute risk	95% CI	P-value	Relative absolute risk	95% CI	P-value
disease:ALL	–	–	–	–	–	–
disease:AML	0.84	[0.66; 1.06]	0.15	0.97	[0.77; 1.22]	0.8
disease:CML	0.59	[0.45; 0.78]	0.00023	1.4	[1.13; 1.62]	0.0009
Karnofsky	1.3	[l.01; 1.65]	0.041	0.81	[0.7; 0.94]	0.0056
donor:sibling	–	–	–	–	–	–
matched/unrelated	0.76	[0.57; 1]	0.051	1.7	[1.45; 1.96]	< 0.0001
mismatched/unrelated	0.27	[0.13; 0.57]	0.00061	2.2	[1.85; 2.63]	< 0.0001
stage:early	–	–	–	–	–	–
stage:intermediate	1.8	[1.33; 2.39]	0.00012	1.2	[1.04; 1.42]	0.015
stage:advanced	3.2	[2.51; 4.09]	< 0.0001	1.3	[1.08; 1.59]	0.0056
transplant waiting time	0.99	[0.98; 1]	0.021	1	[1; 1.01]	0.00010

Table 3

Results for donor effects from different competing risk regression models.

Donor	*Relapse*			*Death in remission*		
	exp(β)	**95% CI**	**P-value**	**exp(β)**	**95% CI**	**P-value**
	Absolute risk regression					
matched/unrelated vs. sibling	0.76	[0.57; 1]	0.051	1.7	[1.45; 1.96]	< 0.0001
mismatched/unrelated vs. sibling	0.27	[0.13; 0.57]	0.00061	2.2	[1.85; 2.63]	< 0.0001
	Logistic-link regression					
matched/unrelated vs. sibling	0.65	[0.43; 0.98]	0.038	2.3	[1.74; 3]	< 0.0001
mismatched/unrelated vs. sibling	0.16	[0.06; 0.41]	0.00013	5	[3.21; 7.89]	< 0.0001
	Fine-Gray regression					
matched/unrelated vs. sibling	0.7	[0.5; 0.99]	0.042	1.9	[1.57; 2.37]	< 0.0001
mismatched/unrelated vs. sibling	0.21	[0.09; 0.48]	0.00025	3.3	[2.46; 4.41]	< 0.0001
	Cause-specific Cox regression					
matched/unrelated vs. sibling	1.05	[0.78; 1.42]	0.75	2.08	[1.71; 2.52]	< 0.0001
mismatched/unrelated vs. sibling	0.40	[0.20; 0.81]	0.01	2.87	[2.20; 3.75]	< 0.0001