Radiotherapy and Oncology 130 (2019) 185–189

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original article

Competing risks in survival data analysis

Almut Dutz [a,b], Steffen Löck [a,c,d,*]

[a] OncoRay – National Center for Radiation Research in Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Helmholtz-Zentrum Dresden - Rossendorf; [b] Helmholtz-Zentrum Dresden - Rossendorf, Institute of Radiooncology – OncoRay; [c] German Cancer Consortium (DKTK), Partner Site Dresden, and German Cancer Research Center (DKFZ), Heidelberg; and [d] Department of Radiotherapy and Radiation Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

ARTICLE INFO

Article history:
Received 7 June 2018
Received in revised form 11 September 2018
Accepted 11 September 2018
Available online 9 October 2018

Keywords:
Competing risk
Survival data
Time-to-event data
Cox regression

ABSTRACT

Clinical trials and retrospective studies in the field of radiation oncology often consider time-to-event data as their primary endpoint. Such studies are susceptible to competing risks, i.e. competing events may preclude the occurrence of the event of interest or modify the chance that the primary endpoint occurs. Competing risks are frequently neglected and the event of interest is analysed with standard statistical methods. Here, we would like to create awareness of the problem and demonstrate different methods for survival data analysis in the presence of competing risks.

Clinical trials and retrospective studies in the field of radiation oncology often consider time-to-event data as their primary endpoint. Here, the period of time until an event of interest occurs is recorded [1,2]. If the study ends before the event has occurred or the patient is no longer available for follow-up, the outcome is typically right-censored at the time of last follow-up. Analyses of such data are commonly performed using (i) the Kaplan–Meier (KM) method to estimate and visualise the probability of survival, (ii) the log-rank test to compare survival between stratified patient groups and (iii) Cox proportional hazards regression to estimate the effect of covariates on survival. These methods assume that censoring is random and non-informative, i.e. censored patients would have the same probability for experiencing the event as the patients still at risk, if they received longer follow-up [3–8]. This assumption might well hold for true survival endpoints like overall survival. For other endpoints, such as local tumour control or the occurrence of late side effects, however, the assumption of random and non-informative censoring does not hold [9]. In these cases, other events – so-called competing risks (CR) – may preclude the occurrence of the event of interest or modify the chance that the primary endpoint occurs [2–6,10–12].

As an example, consider loco-regional relapse (LRR) as the event of interest in a study on patients with head and neck squamous cell carcinoma (HNSCC). Patients that died before experiencing LRR were subject to the competing event of death and cannot experience LRR any longer, i.e. their risk of LRR is zero. In a Kaplan–Meier analysis of LRR, however, patients who died without LRR are censored. This erroneously assumes that in the future they would have the same risk of relapse as the surviving patients, i.e. their risk of LRR is larger than zero. Thus, a substantial overestimation of the incidence of LRR may occur using the Kaplan–Meier approach. This is also important for the comparison of different patient subgroups. For example, a study on HNSCC may investigate the differences in LRR between patients with Human papillomavirus (HPV) positive and HPV-negative tumours. As patients with HPV-negative tumours are often older and consume alcohol and tobacco more regularly than patients with HPV-positive tumours, they are more likely to die (e.g. due to cardiovascular events, second malignancies), i.e. the HPV-negative group has a higher incidence of the competing event [13]. This may lead to a substantial overestimation of the incidence of LRR using the Kaplan–Meier approach in the HPV-negative group. In the HPV-positive group the competing event occurs less, so that the overestimation may be considerably lower. Thus, differences in the incidence of LRR between both groups may in fact be smaller than they appear in a Kaplan–Meier analysis. Hence, appropriate methods accounting for the CR should be applied.

Competing risks are present in many medical articles dealing with survival analysis [5,14,15]: about half of the Kaplan–Meier analyses in medical journals are susceptible to CR [5,11,15,16]. The issue may become even more relevant in the future, e.g. for

* Corresponding author at: OncoRay – National Center for Radiation Research in Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Fetscherstr. 74, 01307 Dresden, Germany.
E-mail address: steffen.loeck@oncoray.de (S. Löck).

https://doi.org/10.1016/j.radonc.2018.09.007
0167-8140/

elderly patients [14] who are more likely to experience several potential disease endpoints, i.e. the occurrence of competing events increases [2,17].

The present manuscript attempts to highlight different methods of survival analysis in the presence of CR. Advantages and limitations of different approaches are addressed. We illustrate CR analyses using an example patient cohort that underwent radiochemotherapy (RCTx) of HNSCC.

Example patient cohort

To demonstrate the CR analyses in more detail we used data from patients with HNSCC who underwent contemporary primary RCTx within the German Cancer Consortium Radiation Oncology Group (DKTK-ROG). The cohort is described elsewhere [18,19]. From this cohort 149 patients with known p16 status were included in the analysis (p16 negative: 125 (84%), p16 positive: 24 (16%)). LRR was chosen as the event of interest. Thus, the CR was death without LRR. p16 status was used for patient classification and as input variable for prognostic modelling. To enhance the impact of the CR on the presented analyses, patients censored before 36 months of complete follow-up were considered to have died without LRR (13 patients out of 42 with the CR). The possible patient states are depicted in Fig. 1.

An extended hypothetic example on twelve patients in the style of the analysis of Kim et al. [21] can be found in the supplement (Appendix A). It includes detailed calculations of the cumulative incidence to compare the conventional KM method with an approach considering CR. A compilation of common terms regarding CR analysis is given in the supplement (Appendix B).

Methods for competing-risk analyses

Estimation of the cumulative incidence

The Kaplan–Meier method is often applied to estimate the cumulative incidence of an event, using the 1-KM estimator (1-KM). This method is appropriate for endpoints such as overall survival, but also for composite endpoints such as progression-free survival. For progression-free survival, both progress and death of any cause are considered as events. In contrast, for the endpoint freedom from progression, death would be a competing event. Thus, composite endpoints may avoid potential complications due to CR. However, they may be more difficult to interpret, especially when combining events with different severities and different incidence rates. It may not be clear to which of these events a possible difference in the composite endpoint between patient subgroups can be attributed.

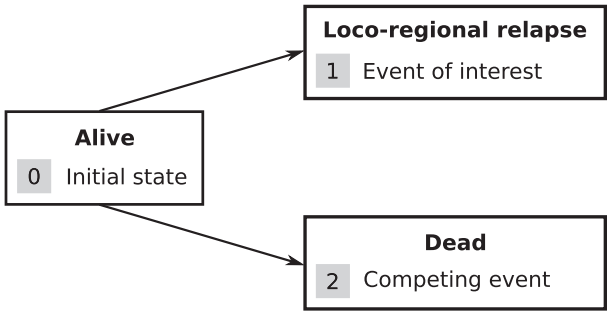

Fig. 1. Potential states of the patients considered for the modified HNSCC cohort. In the initial state, the patient is alive and event-free (0). The first event may be loco-regional relapse (event of interest, 1) or death before the patient experienced loco-regional relapse (competing risk, 2).

Other endpoints are cause-specific, i.e. they consider only a subset of several possible reasons for an event to occur, e.g. cancer-related death. The unconsidered reasons may act as competing events. Such endpoints, which are susceptible to CR, should not be estimated by the Kaplan–Meier approach, since the competing events would be censored as if these patients were lost to follow-up [7]. As an example, Fig. 2 depicts the 1-KM curves for LRR (event of interest, A and B) and death without LRR (competing event, C and D) by dashed lines for the modified HNSCC cohort. In the total absence of the CR, the event of interest would have a greater chance to occur compared to reality, in which the CR may preclude the event of interest from occurring. Thus, the risk of the event of interest estimated by the 1-KM method exceeds its true risk [4–6,8,11,14,22,23]. This overestimation, also called "competing risk bias", may even be as large as the underlying difference between patient groups when comparing different subgroups and is amplified with increasing incidence of the CR [8,22,23]. Consequently, the incidences of all event types calculated by the 1-KM approach may add up to more than 100%, which is illustrated in the hypothetic example in the supplement (Appendix A, Fig. S1B). Due to the overestimation of the endpoint, e.g. LRR, one may be inclined to intensify a therapy resulting in overtreatment.

To estimate the cumulative incidence in the presence of CR, the 1-KM method should be replaced by a method accounting for the competing events, the so-called cumulative incidence function (CIF) [4]. In contrast to the 1-KM approach, the encoding of the patient status is no longer binary (0: censored, 1: event) but contains more categories: 0 for non-informative censoring, 1 for the event of interest and 2, 3, ... for each competing event. As the CIF of any first event is the sum of all CIFs of all event types, the sum of mutually exclusive events will exactly add to 100%. For mathematical details on the 1-KM and the CIF method, see Appendix A. Fig. 2 shows the CIF curves for LRR (event of interest, A and B) and death (competing event, C and D) by solid lines for the modified HNSCC cohort. The overestimation of the 1-KM approach is clearly visible. Even though the steps of both functions occur at the same failure time, their heights differ as explained in Appendix A.

Comparison between patient groups

To evaluate differences in the survival curves between subgroups, the log-rank test is frequently employed [25]. In the presence of CR, the log-rank test would compare the subgroups based on the 1-KM estimates of the endpoint, which is inappropriate as discussed above. To evaluate differences between treatment groups in the presence of CR a test developed by Gray (Gray's test) is widely used [24]. As both tests treat patients suffering from a competing event in different ways, their results may differ.

We illustrate the difference between both tests using the modified HNSCC cohort to test whether patients with p16-positive and p16-negative tumours differ with respect to the occurrence of LRR. Fig. 2 (B, D) shows the cumulative incidences calculated by the 1-KM approach and the CIF for LRR and death without LRR stratified by the p16 status. The subgroups were compared with the respective test. For LRR, the p-values of the log-rank test and Gray's test differ slightly (0.014 and 0.022, respectively) and to a larger extend for the CR: $p = 0.17$ for the log-rank test and $p = 0.64$ for Gray's test. This difference is caused by the higher incidence of LRR in patients with p16-negative tumours. Censoring the CR in the 1-KM analysis leads to a 45% incidence of death after 5 years, while using the CIF reduces the incidence to only 29%. In the patient group with p16-positive tumours the incidences of the CR are similar for both methods (34% and 29% at 5 years, respectively). Thus, the difference between both patient groups in death

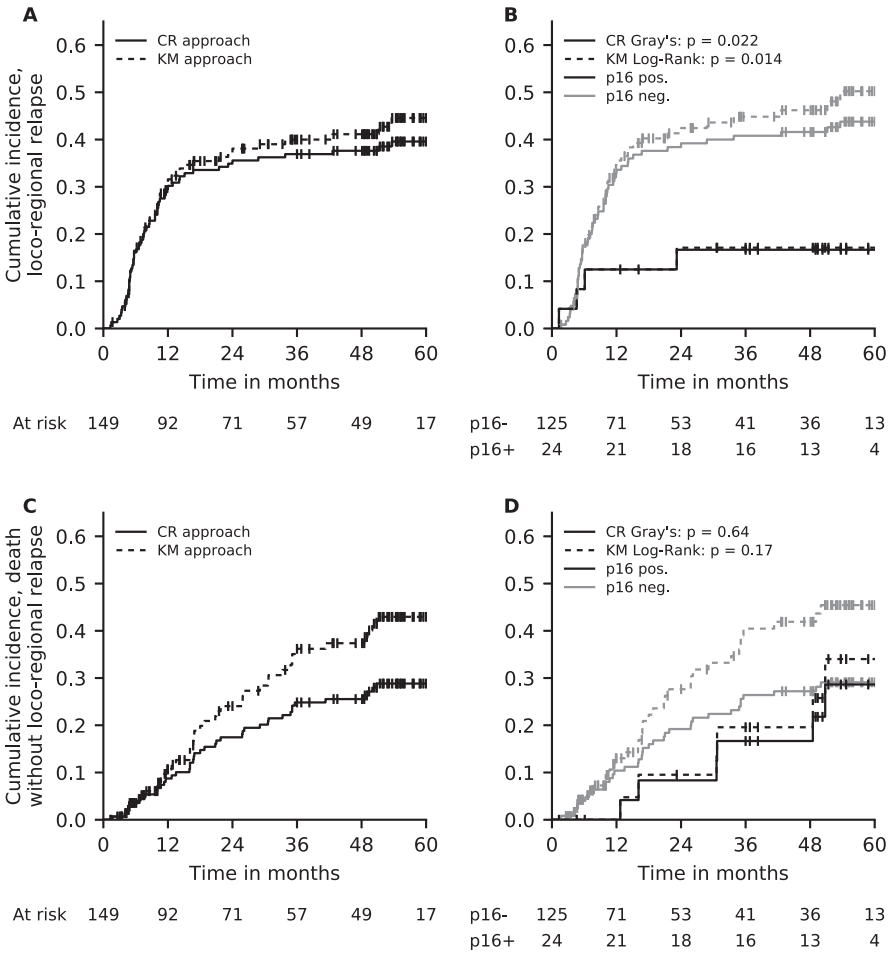

Fig. 2. Comparison of the cumulative incidence function (solid lines) to the Kaplan–Meier approach (dashed lines) for the event of interest loco-regional relapse (A and B) and the competing event death without loco-regional relapse (C and D) using the entire modified HNSCC cohort (left) and stratified for p16 status (right).

is smaller for the CIF (compare the difference between dashed lines to the difference between solid lines in Fig. 2D), explaining the larger *p*-value of Gray's test compared to the log-rank test.

Regression analysis

To assess the association between covariates and the event of interest, a Cox proportional (cause-specific) hazards model (Cox model) is often used in survival analysis. This model is easy to fit and implemented in common software solutions (Appendix C). In the presence of CR, however, the disadvantage of this approach is that the regression coefficients cannot be interpreted directly. They can be understood as the effect of the covariate on the relative increase in the incidence rate of the event of interest among those patients who have not experienced any event yet [22].

In the presence of CR, an estimate of the effect of covariates on the event of interest can be obtained using the Fine–Gray model [26], also referred to as the subdistribution hazard model [6]. The Fine–Gray model returns accurate estimates of the probability of the events of interest that will be seen over time making it more interpretation friendly than the Cox model [9]. Table 1 shows the results of the Cox proportional hazards model and the Fine–Gray model for LRR and for death without LRR in the modified HNSCC cohort. For these univariable models, p16 status was chosen as a potential prognostic variable. The effect of the p16 status on the endpoint is estimated by the cause-specific hazard ratio (HR) or the subdistribution hazard ratio (sHR).

For the Cox model, HR = 0.30 means that in a hypothetic world without any LRR-free deaths, patients with p16-positive tumours are 0.30 times less likely to experience LRR than patients with p16-negative tumours. A sHR of 0.33 for LRR means that in the presence of LRR-free deaths, patients with p16-positive tumours are 0.33 times less likely to have LRR. Since patients with p16-negative tumours showed more LRR-free deaths, they had less chance to experience LRR. This is accounted for in the Fine–Gray model leading to the slightly weaker sHR. The sHR of 0.81 for the competing event death is also weaker than the cause-specific HR estimate of 0.55. Analogously, for patients with p16-negative tumours more relapses were observed such that less LRR-free deaths could occur among them.

In the presented example, the effect of p16 status was in the same direction for the event of interest and for the CR, i.e. patients with p16-positive tumours had a lower incidence for both LRR and death compared to p16-negative patients. In other situations, the effect may be in the opposite direction, i.e. one patient group has a higher incidence for LRR, but lower values for relapse-free deaths compared to a second patient group. Because of the beneficial effect of a low mortality rate in the first group, more patients are available for the event of interest LRR. To uncover such effects it is important to investigate not only the event of interest but also the competing event.

For the sake of simplicity, we limited the analysis to one CR and univariable regression models only. In the case of various independent CR, they can be grouped to one category "CR events" and be

Table 1
Regression results for loco-regional relapse (LRR) and the competing event death without LRR using p16 status as model predictor on the modified HNSCC cohort. CI, confidence interval; HR, hazard ratio; sHR, subdistribution hazard ratio.

	Number of events		Cox model			Fine-Gray model		
Event type	p16 positive (n = 24)	p16 negative (n = 125)	HR	95% CI	p-value	sHR	95% CI	p-value
LRR	4	55	0.30	(0.11–0.83)	0.020	0.33	(0.12–0.94)	0.039
Death without LRR	6	36	0.55	(0.23–1.32)	0.18	0.81	(0.36–1.84)	0.62

Table 2
Time-to-event analyses in the absence or presence of competing risks. CR, competing risk [21].

		Absence of CR	Presence of CR
(i)	Estimating cumulative incidence	Kaplan–Meier method	Cumulative incidence function
(ii)	Comparison of hazards		
	Test to compare hazards for different treatments/groups	Log-rank test	Gray's test [24]
(iii)	Regression analysis		
	Identify important potential prognostic factors	Proportional hazards model (Cox model)	Subdistribution hazards model (e.g. Fine-Gray model [26]) Cause-specific proportional hazards model (for aetiological research questions)

analysed analogously [11]. Regression analysis can be conducted with multiple covariates in the same way [4,20].

Table 2 summarises the presented methods appropriate for survival analyses in the presence of CR and contrasts them with the well-known standard methods.

Conclusion

In the presence of competing risks, *Radiotherapy and Oncology* may in the future ask authors to perform competing risk analyses as outlined in this manuscript. The following points may be considered:

(1) Be aware that the endpoint of interest may be susceptible for CR [21].
(2) Figure out which competing events may be experienced by the study population and may change the probability of the occurrence of the event of interest (Appendix E) [11].
(3) Avoid the conventional Kaplan–Meier approach, whenever competing risks are present [3,5–8,11,14,15,23]. In the absence of CR, the standard Kaplan–Meier approach, log-rank test and Cox proportional hazards model are appropriate (e.g. for overall or progression-free survival) [27].
(4) Calculate the cumulative incidence function of an event of interest in the presence of CR, perform a proper test for comparison of hazards between patient groups (Gray's test) and include CR regression analyses [21]. An introductory and an advanced guide for regression modelling in CR settings using R are presented by Scrucca et al. [10,28].
(5) Present the results for all event types, also for the competing events [21].
(6) Interpret the results in collaboration between statisticians and clinicians [11].
(7) If it is not possible to adjust for CR, discuss CR as an important study limitation [12].

Note that both, the cause-specific Cox proportional hazards regression and the Fine–Gray subdistribution hazards regression model may provide particular insights on the effect of the covariates on the endpoint of interest. Hence, a full understanding may require an analysis using both methods in some settings [6,20,23]. Overall, there can be no general recommendation for all problems. The choice of an appropriate method depends on the question of interest and on the incidence rate of the CR [17,27,29]. Besides the analysis of tumour-related outcome, clinical studies may also investigate severe late normal-tissue effects following RT. A discussion about potential difficulties and pitfalls investigating late side effects can be found in Appendix D.

Readers should keep in mind, that Kaplan–Meier estimates are biased upward whenever competing events are present [5,12]. Moreover, one should be careful to apply results of cause-specific hazard models from published studies to populations with other incidence rates of the competing events.

Acknowledgments

The authors would like to thank Matthias Kuhn for his constructive criticism of the manuscript, and the German Cancer Consortium Radiation Oncology Group (DKTK-ROG) for providing the example HNSCC dataset.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2018.09.007.

References

[1] Lee ET, Go OT. Survival analysis in public health research. Annu Rev Public Health 1997;18:105–34.
[2] Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. Br J Cancer 2004;91:1229–35.
[3] Varadhan R, Weiss CO, Segal JB, Wu AW, Scharfstein D, Boyd C. Evaluating health outcomes in the presence of competing risks: a review of statistical methods and clinical applications. Med Care 2010;48:S96–S105.
[4] Noordzij M, Leffondré K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? Nephrol Dial Transplant 2013;28:2670.
[5] van Walraven C, McAlister FA. Competing risk bias was common in Kaplan-Meier risk estimates published in prominent medical journals. J Clin Epidemiol 2016;69:e8.
[6] Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. Circulation 2016;133:601–9.
[7] Wolbers M, Koller MT, Stel VS, et al. Competing risks analyses: objectives and approaches. Eur Heart J 2014;35:2936.
[8] Verduijn M, Grootendorst DC, Dekker FW, Jager KJ, le Cessie S. The analysis of competing events like cause-specific mortality – beware of the Kaplan-Meier method. Nephrol Dial Transplant 2011;26:56.
[9] Caplan RJ, Pajak TF, Cox JD. Analysis of the probability and risk of cause-specific failure. Int J Radiat Oncol Biol Phys 1994;29:1183–6.
[10] Scrucca L, Santucci A, Aversa F. Competing risk analysis using R: an easy guide for clinicians. Bone Marrow Transplant 2007;40:381–7.
[11] Pintilie M. An introduction to competing risks analysis. Rev Esp Cardiol (Engl Ed) 2011;64:599–605.
[12] Wongworawat MD, Dobbs MB, Gebhardt MC, et al. Editorial: estimating survivorship in the face of competing risks. Clin Orthop 2015;473:1173–6.
[13] Dok R, Nuyts S. HPV positive head and neck cancers: molecular pathogenesis and evolving treatment strategies. Cancers 2016;8:41.
[14] Koller MT, Raatz H, Steyerberg EW, Wolbers M. Competing risks and the clinical community: irrelevance or ignorance? Stat Med 2012;31:1089–97.

[15] Schumacher M, Ohneberg K, Beyersmann J. Competing risk bias was common in a prominent medical journal. J Clin Epidemiol 2016;80:135–6.
[16] Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. Stat Med 1999;18:695–706.
[17] Dignam JJ, Zhang Q, Kocherginsky M. The use and interpretation of competing risks regression models. Clin Cancer Res 2012;18:2301–8.
[18] Lohaus F, Linge A, Tinhofer I, et al. HPV16 DNA status is a strong prognosticator of loco-regional control after postoperative radiochemotherapy of locally advanced oropharyngeal carcinoma: Results from a multicentre explorative study of the German Cancer Consortium Radiation Oncology Group (DKTK-ROG). Radiother Oncol 2014;113:317–23.
[19] Linge A, Lohaus F, Löck S, et al. HPV status, cancer stem cell marker expression, hypoxia gene signatures and tumour volume identify good prognosis subgroups in patients with HNSCC after primary radiochemotherapy: A multicentre retrospective study of the German Cancer Consortium Radiation Oncology Group (DKTK-ROG). Radiother Oncol 2016;121:364–73.
[20] Latouche A, Allignol A, Beyersmann J, Labopin M, Fine JP. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. J Clin Epidemiol 2013;66:648–53.
[21] Kim HT. Cumulative incidence in competing risks data and competing risks regression analysis. Clin Cancer Res 2007;13:559–65.
[22] Wolbers M, Koller MT, Witteman JCM, Steyerberg EW. Prognostic models with competing risks: methods and application to coronary risk prediction. Epidemiology 2009;20:555–61.
[23] Wolkewitz M, Cooper BS, Bonten MJM, Barnett AG, Schumacher M. Interpreting and comparing risks in the presence of competing events. Br Med J 2014;349.
[24] Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat 1988;16:1141–54.
[25] Mantel N. Evaluation of survival data and two rank order statistics in its consideration. Cancer Chemother Rep 1966;50:163–70.
[26] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496–509.
[27] Chappell R. Competing risk analyses: how are they different and why should you care? Clin Cancer Res 2012;18:2127–9.
[28] Scrucca L, Santucci A, Aversa F. Regression modeling of competing risk using R: an in depth guide for clinicians. Bone Marrow Transplant 2010;45:1388.
[29] Dignam JJ, Kocherginsky MN. Choice and interpretation of statistical tests used when competing risks are present. J Clin Oncol 2008;26:4027–34.