COMPARISONS BETWEEN THE KAPLAN-MEIER COMPLEMENT AND THE CUMULATIVE INCIDENCE FOR SURVIVAL PREDICTION IN THE PRESENCE OF COMPETING EVENTS

by

Jeffrey D. Borrebach

B.A., Case Western Reserve University, 2010

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Jeffrey D. Borrebach

It was defended on

April 15, 2013

and approved by

Jong-Hyeon Jeong, Ph.D., Assistant Professor, Biostatistics, Graduate School of Public Health, University of Pittsburgh

Maria Mori Brooks, Ph.D., Associate Professor, Epidemiology, Graduate School of Public Health, University of Pittsburgh

Thesis Advisor: Stewart J. Anderson, Ph.D., Professor, Biostatistics, Graduate School of Public Health, University of Pittsburgh

Stewart J. Anderson, Ph.D.

COMPARISONS BETWEEN THE KAPLAN-MEIER COMPLEMENT AND THE CUMULATIVE INCIDENCE FOR SURVIVAL PREDICTION IN THE PRESENCE OF COMPETING EVENTS

Jeffrey D. Borrebach, M.S.

University of Pittsburgh, 2013

ABSTRACT

Estimating cumulative event probabilities in time-to-event data can be complicated by competing events. Competing events occur when individuals can experience events other than the primary event of interest. These "other events" are often treated as censored observations.

This thesis compares point estimates and relative differences between two cumulative event probability estimators, the Kaplan-Meier complement (KMC) and the cumulative incidence (CI), in the presence of competing events. The KMC does not allow for the possibility of experiencing competing events, whereas the CI does. Consequently, the KMC overestimates the CI in the presence of competing events.

In this thesis, data were simulated with different combinations of primary event hazards, competing event hazards, random censoring hazards, and sample sizes. Cumulative event probabilities using the KMC and CI methods were calculated over a time period of 10 years.

Several conclusions were drawn. High primary event hazards resulted in high KMC's and CI's and slightly narrowed the variability of the relative differences between the two estimates. High competing event hazards did not affect KMC's but resulted in low CI's, causing high

relative differences. High random censoring hazards did not affect KMC's, CI's, or relative differences. Large sample sizes did not affect the median relative differences but did narrow the variability of the relative differences.

The public health relevance of this thesis is to help medical clinicians and researchers understand the advantages and disadvantages of different approaches of calculating cumulative event probabilities in situations where competing events occur. This is particularly important in the area of personalized medicine in diseases like cancer where clinicians attempt to predict their patients' mortality or recurrence probabilities over time given certain clinical, pathologic, or demographic characteristics.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would primarily like to thank my advisor, Dr. Stewart Anderson, for all the help he has given me during the work on this thesis. Second, I would like to thank Dr. Jong-Hyeon Jeong and Dr. Maria Brooks for serving as my thesis committee members. Thanks to Dr. Jong-Hyeon Jeong also for exposing me to survival analysis and suggesting the data simulation algorithm outlined in Section 3.1. Finally, I would like to thank my family for all their love and support throughout my time in graduate school.

1.0 INTRODUCTION

Survival analysis involves a group of techniques in statistics used to analyze time-to-event data, or data involving the time to an occurrence of some event. Events may be as varied as death, the development of a disease, recurrence of a disease, or cessation of a disease.[9] In cases where an event may be experienced multiple times, one is often interested in the first occurrence of that event. The individual who experienced the event is then precluded from the risk set at the time of the first occurrence. Survival analysis has historically been applied to biomedical research, but it also has applications in computer science, engineering, business, economics, criminology, and other fields.[9]

1.1 CENSORING AND COMPETING EVENTS

Survival analysis is complicated by the issue of censoring. Censoring arises when an individual's response or survival length is only known in a certain period of time. The two most common types of censoring are left censoring and right censoring. Left censoring occurs when the event of interest lies before a certain time, but it is unknown by how much.[9] This is typically not a problem in prospective studies where the initiation of accrual is often fixed and the nature of the treatment does not change over calendar time. On the other hand, right censoring, which is

usually more problematic, occurs when the event of interest has not been observed at the last available time, and hence, it is unknown when or even if the event would have occurred.[9]

Censored observations may not only be due to losses to follow-up or administrative cessation of the time period of consideration but can also be due to events not of interest. This situation is problematic if these "other events" preclude observation of the primary event under consideration. For example, in a study analyzing long-term survival rates of cancer patients, a distant metastasis might "censor" a locoregional recurrence in breast cancer patients.[10] These secondary censoring events not due to loss of follow-up or administrative cessation are referred to as *competing events*.[1–10]

The issue of how to deal with competing events is important in survival analysis. Experiencing a competing event acts as a right censor on the primary event. Because of this extra censoring, it is often useful to estimate and compare cumulative event probabilities of a specific event, rather than of all events as a whole.

In this thesis, we propose to consider this problem by comparing the differences between two cumulative event probability estimators in the presence of a competing event using simulations. Section 1 introduces survival analysis and competing events. Section 2 defines the cumulative event probability estimators under consideration. Section 3 contains the methods. Section 4 presents the results. Section 5 discusses the results and contains suggestions to clinical researchers.

2.0 ESTIMATORS OF CUMULATIVE EVENT PROBABILITY

2.1 KAPLAN-MEIER COMPLEMENT

The first event-specific cumulative probability estimator under consideration is the Kaplan-Meier complement. For event times $T_1,\ldots,T_j$, the *event-specific hazard* for event i at time t, $\lambda_i(t)$, is defined as the rate of event i within the next instant of time, $t + \Delta t$:

$$\lambda_i(t) = \lim_{\Delta t \to 0} \frac{\Pr\{t \leq T < t + \Delta t, I = i \mid T \geq t\}}{\Delta t}, i = 1, \ldots, k \qquad (1)$$

where I is the event type indicator and $T = min(T_1,\ldots,T_j)$.[3–5, 8, 9] The *event-specific cumulative hazard* for event i at time t, $\Lambda_i(t)$, is defined as the sum of the event-specific hazards belonging to event i up to time t:[5, 9]

$$\Lambda_i(t) = \int_0^t \lambda_i(u)\Delta u\,, i = 1, \ldots, k. \qquad (2)$$

The *Kaplan-Meier complement* (KMC) of event i at time t is defined as the cumulative probability of experiencing event i before time t in the absence of competing events:

$$KMC_i(t) = \int_0^t \lambda_i(u)S_i(u)\Delta u = 1 - S_i(t) = 1 - \exp\big(-\Lambda_i(t)\big)\,, i = 1, \ldots, k \qquad (3)$$

where the *event-specific survival function* for event i, $S_i(t)$, is defined as the probability that $T \geq t$ when $I = i$.[4-6] $S_i(t)$ can be estimated by the *Kaplan-Meier estimator*:

$$S_i(t) = \prod_{j=1}^{s}\left(1 - \frac{d_{i_j}}{n_{i_{j-1}}}\right), i = 1, \ldots, k \quad (4)$$

where *s* is the largest *j* such that $t_j < t$, d_{i_j} is the number of individuals who experience event *i* at time *j*, and n_{i_j} is the number of individuals at risk for event *i* beyond time t_j.[5, 6, 9]

It is important to note that this estimator is interpretable only if events due to all other causes are removed. The probability of experiencing a competing event prior to time *t* is assumed to be zero when this does not actually reflect the true situation under competing events. Thus, the KMC cannot be considered the true probability of event *i* occurring before time *t* because competing events are treated as censored observations.[2, 5, 6] A presumably more appropriate estimator of cumulative event probability, the cumulative incidence, correctly allows for the possibility of experiencing competing events.

2.2 CUMULATIVE INCIDENCE

The other event-specific cumulative probability estimator under consideration is the cumulative incidence. The *cumulative incidence* (CI) of event *i* at time *t* is defined as the cumulative probability of experiencing event *i* before time *t* in the presence of competing events:[2–7, 9]

$$CI_i(t) = \int_0^t \lambda_i(u)S(u)\Delta u\,, i = 1, \ldots, k \quad (5)$$

where the *overall survival function, S(t),* is defined as the probability that $T \geq t$:[4, 5, 8, 9]

$$S(t) = \exp\left(-\sum_{i=1}^{k}\Lambda_i(t)\right) = \prod_{i=1}^{k} S_i(t)\,, i = 1, \ldots, k. \quad (6)$$

The CI can alternatively be derived from the subdistribution hazard instead of the event-specific hazard. The subdistribution hazard for event *i* represents the hazard function for an artificial time variable *T'*, defined as *T'* = *T* if an event of type *i* occurred and *T'* = ∞ if an event from a competing event occurred.[3, 7]

$CI_i(t)$ can be estimated similarly to (4):

$$CI_i(t) = \sum_{j=1}^{s} \frac{d_{i_j}}{n_{i_{j-1}}} S(t_j), i = 1, \dots, k \qquad (7)$$

where *s* is the largest *j* such that $t_j < t$, d_{i_j} is the number of individuals who experience event *i* at time *j*, and n_{i_j} is the number of individuals at risk for event *i* beyond time t_j.[6]

The CI gives a more accurate representation of the cumulative event probability than the KMC in the presence of competing events because competing events are included in the risk set (the KMC is equivalent to the CI in the absence of competing events).[5, 6] The CI does have one disadvantage, though. Unlike the KMC, individuals who experience a competing event before time *t* are still considered to be in the CI risk set for event *i* at time *t*.[3] These individuals are not at risk at time *t* in reality. Because of this and other reasons to be discussed, clinicians and researchers examining cumulative event probabilities may want to use the KMC rather than the CI.

The KMC always overestimates the CI in the presence of competing events because reducing the number of individuals in the risk set inflates the proportion of individuals at risk.[5, 6] Central to this thesis is analyzing the effects of changing various underlying parameters on the differences between these two estimators.

2.3 RELATIVE DIFFERENCE BETWEEN THE KAPLAN-MEIER COMPLEMENT AND THE CUMULATIVE INCIDENCE

Because the KMC overestimates the CI, it is useful to devise a metric that conveys the difference between the two estimators clearly to investigators. The *absolute difference*, $KMC_i(t) - CI_i(t)$, is the simplest metric to use, but it has several disadvantages. As time increases, the difference between KMC and CI monotonically increases. Because of this, it is hard to compare absolute differences between earlier and later time points. Also, the absolute difference between KMC and CI generally tends to be smaller when the CI is small and larger when the CI is large. A measure of *relative difference* (RD) between the two estimators nullifies the disparities in time and CI between different models:

$$RD_i(t) = \frac{KMC_i(t) - CI_i(t)}{CI_i(t)} \quad (8)$$

3.0 METHODS

We conducted simulations in the statistical programming language R 2.13.1[11] to compare the estimates of and relative differences between the KMC and CI of a primary event when competing events could occur. The parameters were chosen based on results from a recent paper for which the authors predicted locoregional recurrence (LRR) probabilities in the presence of competing events in a population of stage II breast cancer patients after receiving neoadjuvant chemotherapy (NC).[10] The parameters and their values tested are as follows:

1) Primary event hazard (λ_1): 0.015 and 0.03
2) Competing event hazard (λ_2): 0.015, 0.05, and 0.1
3) Random censoring hazard (λ_C): 0.04 and 0.08
4) Sample size (N): 250 and 1000

There were 24 different parameter and sample size combinations in total. For each combination, 5,000 independent data sets were simulated and the KMC, CI, and RD estimates were calculated every year from years 1–10. Primary, competing, and random censoring event times were randomly generated for each data simulation according to a binomial algorithm developed by Beyersmann et al. (see Section 3.1).[1] The event and random censoring times were assumed to follow the exponential distribution with cumulative distribution function:[9]

$$F_\lambda(x) = 1 - \exp(-\lambda x), x \geq 0. \quad (9)$$

For each simulation, a survival curve was fit to obtain the Kaplan-Meier survival estimates for the primary event (function 'survfit' in R). The KMC's were calculated by subtracting the Kaplan-Meier estimates from 1. A Fine-Gray cumulative incidence model was also fit to obtain the cumulative incidence estimates for the primary event (function 'cuminc' in R).[3] Finally, the RD's were calculated from the KMC's and CI's. In the case that the KMC and/or CI was inestimable at later time points due to having no remaining individuals in the risk set, the maximum estimate was appended to those later time points.

3.1 DATA SIMULATION ALGORITHM

Simulations of competing events data are often conducted using some "latent event time" model, in which event times and types are associated with the minimum time of several possibly unobserved events. The latent event time model has been criticized for lack of plausibility in some settings, and further, non-identifiability of the dependence structure occurs among postulated latent event times.[1] Therefore, we decided to adopt an algorithm developed by Beyersmann et al. (2009) to generate conditional primary and competing event times.[1] Censoring times were assumed to be independent of primary and competing event times. The steps of the algorithm are as follows:

1) The primary event hazard, λ_1, and competing event hazard, λ_2, are predefined (see Section 3.0).
2) Event times are randomly generated from the exponential distribution (9) with rate $\lambda_1 + \lambda_2$.

3) For all event times, a binomial experiment decides with probability $\lambda_1 / (\lambda_1 + \lambda_2)$ on the primary event. All other events are designated as the competing event.
4) Independent censoring times are randomly generated from the exponential distribution (9) with rate λ_C.

4.0 RESULTS

As was indicated in Section 3.0, we conducted 5,000 independent simulations for 24 combinations of primary event hazard, competing event hazard, random censoring hazard, and sample size. Point estimates of the Kaplan-Meier complement (KMC) and cumulative incidence (CI) for the primary event at years 2, 5, and 10 are given in Tables 1 and 2 respectively. Their relative differences (RD) are given in Table 3, and the proportions of censored individuals are given in Table 4. For each result cell in Tables 1–4, the bolded first number represents the median of the 5,000 simulations and the numbers in parentheses represent the 2.5%–97.5% range of the 5,000 simulations. For Figures 1–4, the middle line represents the median estimates, the lower line represents the 2.5% estimates, and the upper line represents the 97.5% estimates.

Table 1. Kaplan-Meier Complement (KMC) estimates

λ_1	λ_2	λ_C	N	Year 2	Year 5	Year 10
0.015	0.015	0.04	250	**0.029** (0.012-0.053)	**0.072** (0.04-0.109)	**0.139** (0.09-0.191)
			1000	**0.029** (0.019-0.041)	**0.072** (0.056-0.09)	**0.139** (0.115-0.164)
		0.08	250	**0.029** (0.009-0.053)	**0.071** (0.037-0.111)	**0.137** (0.086-0.198)
			1000	**0.03** (0.019-0.041)	**0.072** (0.054-0.091)	**0.139** (0.112-0.168)
	0.05	0.04	250	**0.03** (0.009-0.054)	**0.071** (0.038-0.111)	**0.137** (0.086-0.197)
			1000	**0.029** (0.019-0.041)	**0.072** (0.055-0.09)	**0.138** (0.112-0.167)
		0.08	250	**0.028** (0.009-0.054)	**0.071** (0.036-0.113)	**0.138** (0.082-0.203)
			1000	**0.029** (0.019-0.041)	**0.072** (0.053-0.093)	**0.139** (0.109-0.171)
	0.1	0.04	250	**0.028** (0.009-0.055)	**0.071** (0.037-0.113)	**0.138** (0.079-0.207)
			1000	**0.029** (0.019-0.041)	**0.072** (0.053-0.092)	**0.139** (0.107-0.171)
		0.08	250	**0.029** (0.009-0.055)	**0.071** (0.034-0.116)	**0.137** (0.071-0.215)
			1000	**0.029** (0.018-0.041)	**0.072** (0.052-0.093)	**0.139** (0.104-0.176)
0.03	0.015	0.04	250	**0.058** (0.029-0.089)	**0.138** (0.094-0.188)	**0.258** (0.196-0.323)
			1000	**0.058** (0.044-0.074)	**0.139** (0.116-0.162)	**0.259** (0.227-0.291)
		0.08	250	**0.057** (0.03-0.09)	**0.137** (0.092-0.188)	**0.256** (0.186-0.33)
			1000	**0.058** (0.044-0.074)	**0.139** (0.115-0.164)	**0.259** (0.225-0.295)
	0.05	0.04	250	**0.057** (0.03-0.09)	**0.138** (0.093-0.19)	**0.257** (0.19-0.33)
			1000	**0.058** (0.043-0.074)	**0.139** (0.116-0.164)	**0.259** (0.225-0.294)
		0.08	250	**0.058** (0.03-0.092)	**0.138** (0.09-0.194)	**0.258** (0.181-0.341)
			1000	**0.058** (0.043-0.074)	**0.139** (0.115-0.165)	**0.258** (0.22-0.299)
	0.1	0.04	250	**0.057** (0.029-0.092)	**0.138** (0.088-0.194)	**0.258** (0.179-0.341)
			1000	**0.058** (0.043-0.074)	**0.139** (0.114-0.167)	**0.259** (0.219-0.302)
		0.08	250	**0.057** (0.028-0.091)	**0.138** (0.086-0.196)	**0.255** (0.17-0.357)
			1000	**0.058** (0.043-0.074)	**0.139** (0.111-0.166)	**0.258** (0.213-0.306)

Table 2. Cumulative Incidence (CI) estimates

λ_1	λ_2	λ_C	N	Year 2	Year 5	Year 10
0.015	0.015	0.04	250	**0.029** (0.012-0.053)	**0.069** (0.039-0.105)	**0.129** (0.084-0.178)
			1000	**0.029** (0.019-0.04)	**0.07** (0.054-0.087)	**0.129** (0.107-0.152)
		0.08	250	**0.029** (0.009-0.052)	**0.069** (0.036-0.108)	**0.129** (0.08-0.185)
			1000	**0.029** (0.019-0.041)	**0.07** (0.052-0.088)	**0.13** (0.104-0.156)
	0.05	0.04	250	**0.029** (0.008-0.05)	**0.064** (0.034-0.098)	**0.11** (0.068-0.155)
			1000	**0.028** (0.018-0.039)	**0.064** (0.049-0.08)	**0.11** (0.089-0.132)
		0.08	250	**0.027** (0.009-0.052)	**0.063** (0.033-0.1)	**0.11** (0.066-0.159)
			1000	**0.028** (0.018-0.039)	**0.064** (0.048-0.082)	**0.11** (0.087-0.134)
	0.1	0.04	250	**0.025** (0.008-0.05)	**0.056** (0.03-0.089)	**0.088** (0.052-0.129)
			1000	**0.027** (0.017-0.037)	**0.057** (0.043-0.072)	**0.089** (0.07-0.108)
		0.08	250	**0.026** (0.008-0.049)	**0.057** (0.028-0.09)	**0.089** (0.049-0.134)
			1000	**0.027** (0.017-0.037)	**0.057** (0.042-0.073)	**0.089** (0.069-0.111)
0.03	0.015	0.04	250	**0.057** (0.029-0.088)	**0.133** (0.091-0.182)	**0.241** (0.183-0.302)
			1000	**0.057** (0.043-0.073)	**0.134** (0.112-0.157)	**0.241** (0.212-0.272)
		0.08	250	**0.056** (0.03-0.088)	**0.133** (0.089-0.182)	**0.24** (0.174-0.308)
			1000	**0.057** (0.043-0.073)	**0.134** (0.111-0.158)	**0.241** (0.21-0.274)
	0.05	0.04	250	**0.054** (0.029-0.086)	**0.123** (0.083-0.169)	**0.206** (0.152-0.263)
			1000	**0.055** (0.041-0.07)	**0.124** (0.103-0.145)	**0.207** (0.18-0.235)
		0.08	250	**0.055** (0.029-0.087)	**0.123** (0.08-0.173)	**0.206** (0.147-0.271)
			1000	**0.055** (0.041-0.07)	**0.123** (0.102-0.146)	**0.206** (0.177-0.237)
	0.1	0.04	250	**0.053** (0.025-0.083)	**0.11** (0.07-0.153)	**0.167** (0.119-0.22)
			1000	**0.053** (0.039-0.067)	**0.11** (0.09-0.132)	**0.168** (0.143-0.194)
		0.08	250	**0.052** (0.026-0.082)	**0.11** (0.069-0.155)	**0.167** (0.115-0.227)
			1000	**0.053** (0.039-0.067)	**0.11** (0.089-0.132)	**0.168** (0.141-0.196)

Table 3. Relative Difference (RD) estimates

λ_1	λ_2	λ_C	N	Year 2	Year 5	Year 10
0.015	0.015	0.04	250	**0.012** (0.001-0.03)	**0.034** (0.014-0.061)	**0.071** (0.039-0.112)
			1000	**0.014** (0.008-0.022)	**0.037** (0.025-0.05)	**0.073** (0.056-0.094)
		0.08	250	**0.012** (0-0.031)	**0.033** (0.013-0.062)	**0.069** (0.036-0.114)
			1000	**0.014** (0.008-0.023)	**0.037** (0.025-0.05)	**0.074** (0.056-0.093)
	0.05	0.04	250	**0.047** (0.017-0.09)	**0.124** (0.069-0.195)	**0.257** (0.165-0.372)
			1000	**0.05** (0.034-0.068)	**0.127** (0.1-0.159)	**0.261** (0.212-0.316)
		0.08	250	**0.047** (0.016-0.088)	**0.123** (0.068-0.191)	**0.254** (0.156-0.384)
			1000	**0.05** (0.034-0.068)	**0.127** (0.098-0.16)	**0.26** (0.208-0.321)
	0.1	0.04	250	**0.098** (0.039-0.172)	**0.261** (0.152-0.405)	**0.547** (0.331-0.823)
			1000	**0.101** (0.072-0.134)	**0.265** (0.208-0.329)	**0.559** (0.443-0.69)
		0.08	250	**0.098** (0.036-0.175)	**0.256** (0.141-0.4)	**0.538** (0.301-0.858)
			1000	**0.101** (0.073-0.136)	**0.265** (0.205-0.334)	**0.557** (0.434-0.703)
0.03	0.015	0.04	250	**0.012** (0.001-0.029)	**0.034** (0.015-0.06)	**0.07** (0.04-0.107)
			1000	**0.014** (0.008-0.022)	**0.036** (0.026-0.048)	**0.072** (0.056-0.09)
		0.08	250	**0.012** (0.001-0.03)	**0.033** (0.014-0.06)	**0.068** (0.036-0.109)
			1000	**0.014** (0.008-0.022)	**0.036** (0.025-0.049)	**0.072** (0.055-0.091)
	0.05	0.04	250	**0.047** (0.023-0.081)	**0.122** (0.076-0.182)	**0.248** (0.171-0.346)
			1000	**0.049** (0.036-0.065)	**0.125** (0.102-0.152)	**0.253** (0.213-0.298)
		0.08	250	**0.047** (0.022-0.08)	**0.122** (0.075-0.181)	**0.247** (0.166-0.356)
			1000	**0.049** (0.036-0.065)	**0.125** (0.1-0.154)	**0.253** (0.209-0.303)
	0.1	0.04	250	**0.099** (0.055-0.156)	**0.257** (0.173-0.366)	**0.533** (0.361-0.75)
			1000	**0.102** (0.079-0.129)	**0.262** (0.217-0.313)	**0.542** (0.453-0.645)
		0.08	250	**0.098** (0.051-0.158)	**0.255** (0.166-0.366)	**0.526** (0.344-0.767)
			1000	**0.101** (0.077-0.13)	**0.26** (0.212-0.315)	**0.539** (0.441-0.653)

Table 4. Censoring proportions

λ_1	λ_2	λ_C	N	Censoring Proportion
0.015	0.015	0.04	250	**0.572** (0.508-0.632)
			1000	**0.571** (0.541-0.602)
		0.08	250	**0.728** (0.672-0.78)
			1000	**0.727** (0.7-0.754)
	0.05	0.04	250	**0.38** (0.324-0.44)
			1000	**0.381** (0.351-0.411)
		0.08	250	**0.552** (0.488-0.612)
			1000	**0.551** (0.521-0.581)
	0.1	0.04	250	**0.256** (0.204-0.312)
			1000	**0.258** (0.232-0.286)
		0.08	250	**0.412** (0.348-0.472)
			1000	**0.41** (0.378-0.44)
0.03	0.015	0.04	250	**0.472** (0.412-0.532)
			1000	**0.47** (0.439-0.502)
		0.08	250	**0.64** (0.584-0.7)
			1000	**0.64** (0.611-0.669)
	0.05	0.04	250	**0.332** (0.276-0.392)
			1000	**0.333** (0.304-0.363)
		0.08	250	**0.5** (0.436-0.56)
			1000	**0.5** (0.469-0.531)
	0.1	0.04	250	**0.236** (0.184-0.288)
			1000	**0.235** (0.209-0.262)
		0.08	250	**0.38** (0.32-0.444)
			1000	**0.381** (0.351-0.412)

4.1 PRIMARY EVENT HAZARD

Two different primary event hazards were tested: λ_1 = 0.015 and 0.03. Both the KMC and CI estimates at year 10 were much higher when λ_1 = 0.03 compared to λ_1 = 0.015. For example, at year 10 when λ_2 = 0.05, λ_C = 0.08, and N = 1000, the median KMC = 0.139 when λ_1 = 0.015 while the median KMC = 0.258 when λ_1 = 0.03. For the same parameter combination at year 10, the median CI = 0.11 when λ_1 = 0.015 while the median CI = 0.206 when λ_1 = 0.03. It was expected that both cumulative event probability estimates of the primary event would change along with their hazard, so this result is not surprising.

Although the KMC and CI estimates were greatly affected by changes in λ_1, the relative differences between the two were not affected quite as much. For nearly all combinations, the median RD decreased slightly but remained virtually the same. For example, using the same parameter combination as above at year 10, the median RD = 0.26 when λ_1 = 0.015 while the median RD = 0.253 when λ_1 = 0.03. While the median RD's were hardly affected, the variability of the RD's narrowed slightly. For the same parameter combination at year 10, the RD 2.5%–97.5% range was (0.208–0.321) when λ_1 = 0.015 while the range narrowed to (0.209–0.303) when λ_1 = 0.03.

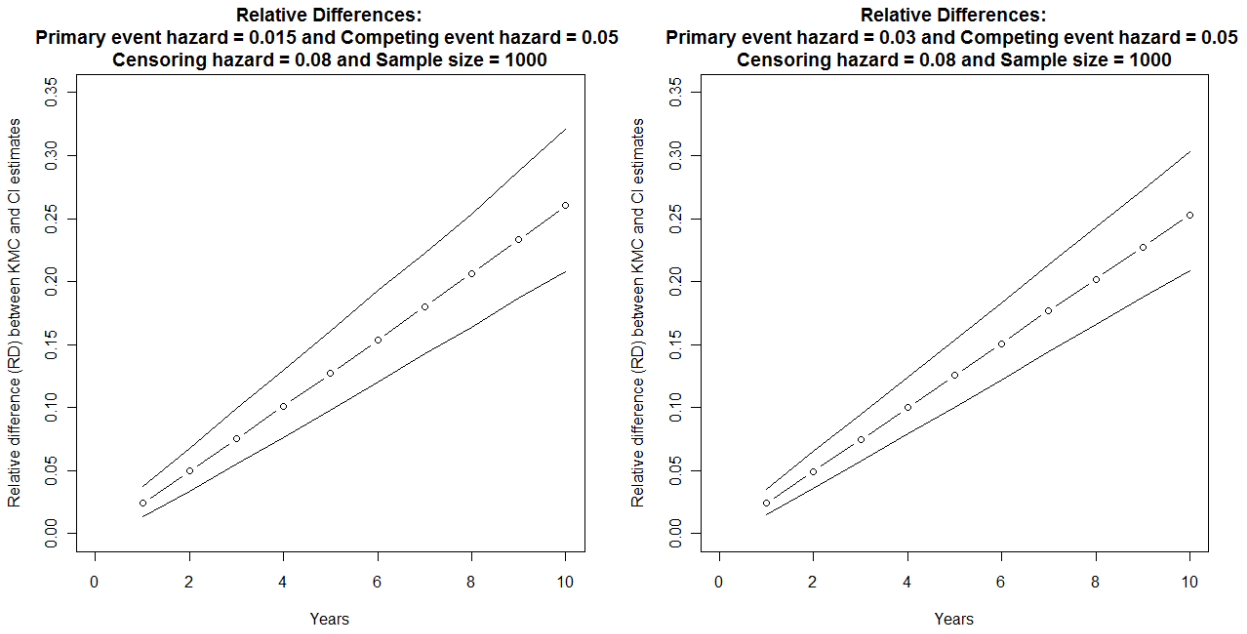

Figure 1. Relative differences between λ_1 = 0.15 and 0.03 (λ_2 = 0.05, λ_C = 0.08, N = 1000)

4.2 COMPETING EVENT HAZARD

Three different competing event hazards were tested: λ_2 = 0.015, 0.05, and 0.1. Higher competing event hazards did not change the KMC estimates. For example, at year 10 when λ_1 = 0.015, λ_C = 0.08, and N = 1000, the median KMC = 0.139 when λ_2 = 0.015 while the median KMC = 0.139 when λ_2 = 0.1. This result was expected, as the calculation of KMC treats competing events as censored observations. Higher competing event hazards did result in reduced CI estimates, however. For the same parameter combination at year 10, the median CI = 0.13 when λ_2 = 0.015 while the median CI = 0.089 when λ_2 = 0.1. This was also expected, as higher rates of competing events reduce the overall survival function (6).

Because the KMC estimates did not change with higher competing hazards but the CI estimates decreased, the RD's dramatically increased. For example, using the same parameter

combination as above at year 10, the median RD = 0.074 when λ_2 = 0.015 while the median RD = 0.557 when λ_2 = 0.1. The variability of the RD's also greatly increased as competing event hazards increased.

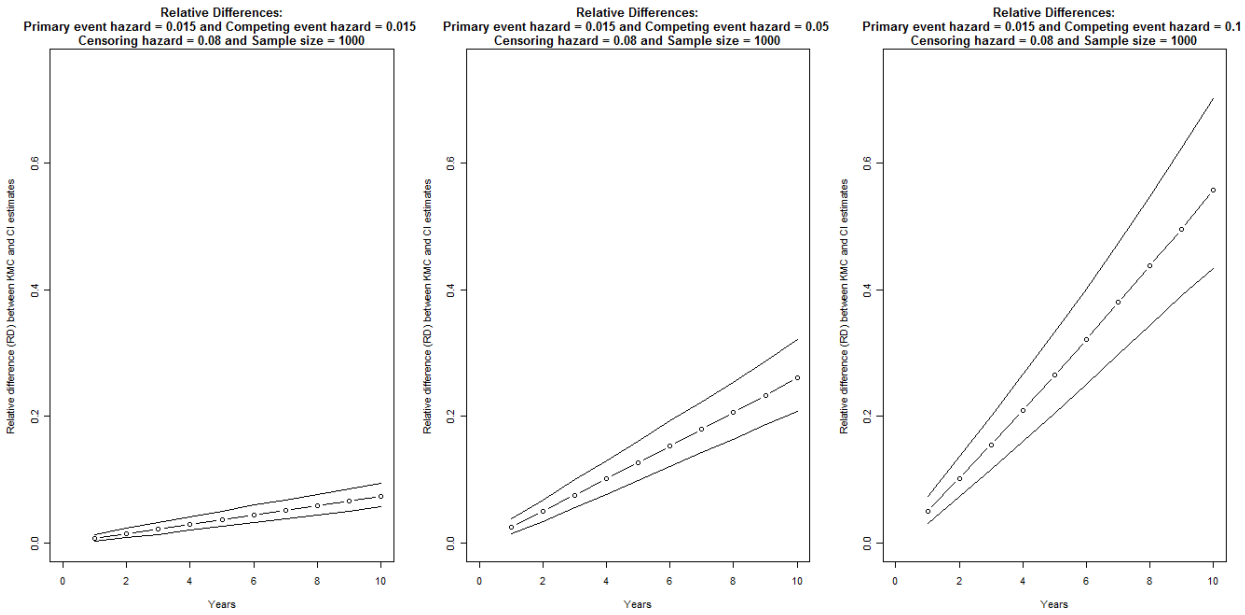

Figure 2. Relative differences between λ_2 = 0.015, 0.05, and 0.1 (λ_1 = 0.015, λ_C = 0.08, N = 1000)

4.3 RANDOM CENSORING HAZARD

Two different random censoring hazards were tested: λ_C = 0.04 and 0.08. As shown in Table 4, median censoring proportions ranged from 0.235 to 0.728. Raising the random censoring hazard from λ_C = 0.04 to λ_C = 0.08 resulted in an increase in the censoring proportion of anywhere from 0.144 to 0.172. Although the censoring proportions increased when the random censoring hazard increased, the KMC and CI estimates did not change. For example, at year 10 when λ_1 = 0.015, λ_2 = 0.05, and N = 1000, the median KMC = 0.138 when λ_C = 0.04

while the median KMC = 0.139 when λ_C = 0.08. For the same parameter combination at year 10, the median CI = 0.11 when λ_C = 0.04 while the median CI = 0.11 when λ_C = 0.08. Because the KMC and CI estimates did not change, the RD's also did not change. For the same parameter combination at year 10, the median RD = 0.261 when λ_C = 0.04 while the median RD = 0.26 when λ_C = 0.08. The variability of the RD's slightly widened, but the effect was so small as to not make any significant difference.

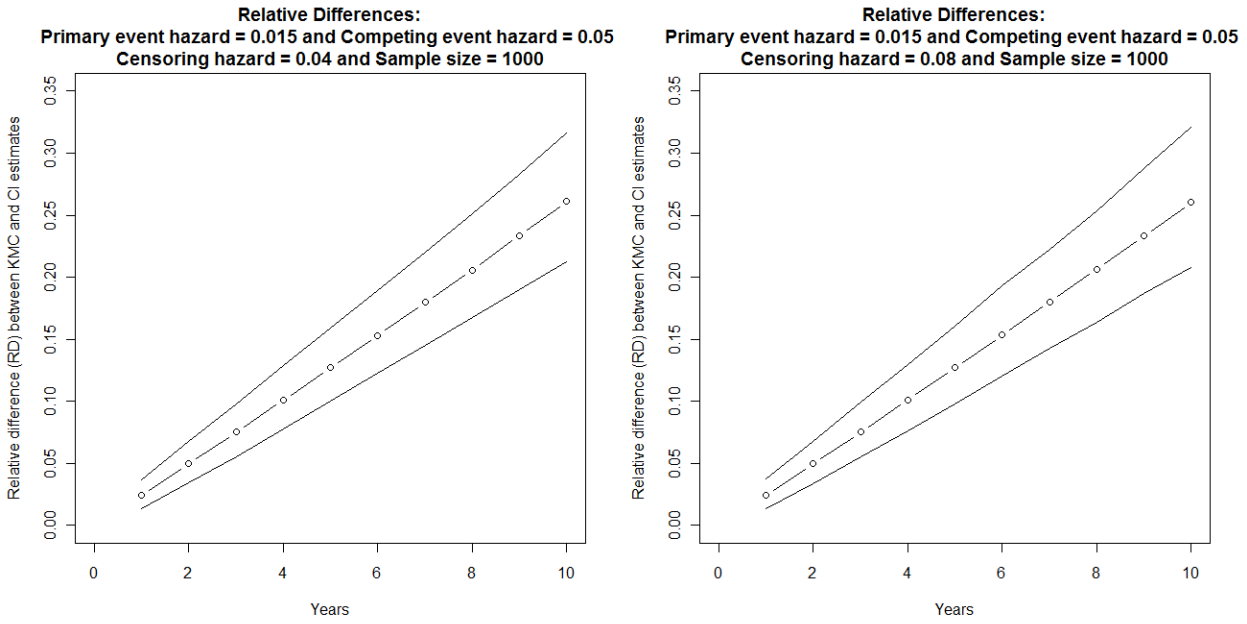

Figure 3. Relative differences between λ_C = 0.04 and 0.08 (λ_1 = 0.015, λ_2 = 0.05, N = 1000)

4.4 SAMPLE SIZE

Two different sample sizes were tested: N = 250 and 1000. Increasing the sample size from 250 to 1000 did not change the median KMC and CI estimates, but it did narrow the variability of those estimates as expected. For example, at year 10 when λ_1 = 0.015, λ_2 = 0.05,

and λ_C = 0.08, the median KMC = 0.138 when *N* = 250 while the median KMC = 0.139 when *N* = 1000. The KMC 2.5%–97.5% range narrowed from (0.082–0.203) when *N* = 250 to (0.109–0.171) when *N* = 1000. Also, for the same parameter combination at year 10, the median CI = 0.11 when *N* = 250 while the median CI = 0.11 when *N* = 1000. The CI 2.5%–97.5% range narrowed from (0.066–0.159) when *N* = 250 to (0.087–0.134) when *N* = 1000.

Because the variability of the KMC and CI estimates narrowed, the variability of the RD's also greatly narrowed. For the same parameter combination as above at year 10, the RD 2.5%–97.5% range narrowed from (0.156–0.384) when *N* = 250 to (0.208–0.321) when *N* = 1000. The median RD's actually slightly increased as sample size increased in most parameter combinations, but the increases were so small as to not make any significant difference (the increases in median RD ranged from only 0.002 to 0.019).

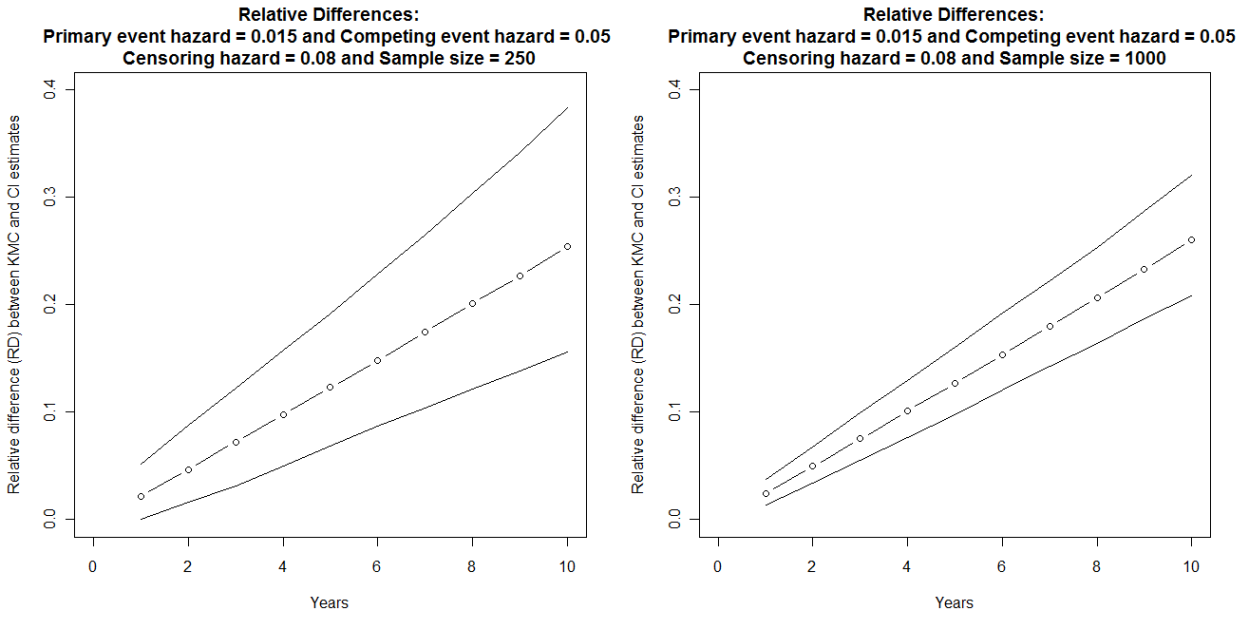

Figure 4. Relative differences between *N* = 250 and 1000 (λ_1 = 0.015, λ_2 = 0.05, λ_C = 0.08)

5.0 DISCUSSION

The issue of whether to use the KMC or the CI as an estimator of cumulative event probability is important in biomedical research. Gooley et al. (1999) unequivocally believe that the CI should always be used as an estimator of cumulative event probability when competing events are present: "Determining whether to use [KMC] or CI is unambiguous and CI should always be used if an estimate of the probability of failure of a specific type is desired."[6] While they are correct that the CI properly takes competing events into account, there still might be reasons for clinicians to want to use the KMC instead.

As an example of KMC's potential clinical advantages, suppose that a woman who has been diagnosed with stage II breast cancer is due to receive a more aggressive treatment if, based on her characteristics, her cumulative event probability is predicted to be above a certain threshold. If her cumulative event probability is predicted to be below that value, she will receive a less aggressive treatment. Suppose also that her KMC estimate lies above this threshold, whereas her CI estimate lies below this threshold. In this case, her clinician may decide to exercise caution and use the KMC, giving the more aggressive treatment and presumably having a greater chance of treating her cancer. Of course, there may also be instances where clinicians would want to avoid overtreatment when the treatments (e.g., certain chemotherapies, radiation therapies, and surgeries used in cancer treatment) have potentially harmful side-effects of their own. In those cases, using the CI may be desired instead. The problem for clinicians becomes

whether they want greater predictive accuracy or to exercise caution in cases where the benefits of overtreatment are perceived to be greater than the risks.

While high primary event hazards resulted in high KMC and CI estimates and vice versa, the RD's were only slightly affected with a small narrowing of their variabilities. The decision to use KMC or CI may not be as much of an issue for clinicians in the case of high primary event hazards, as both the KMC and CI estimates are high. In contrast, high competing event hazards resulted in low CI estimates, leading to dramatically higher RD's. Clinicians who want to err on the side of caution may be compelled to do so if competing event hazards are high. High random censoring hazards were shown to barely affect KMC, CI, or RD estimates, so clinicians should not be concerned about censoring proportions in their studies as long as the censoring is "truly" random and independent of all other events. Finally, as expected, high sample sizes resulted in only a slight change in the median KMC's, CI's, and RD's but a great narrowing of their variabilities. Therefore, clinicians can expect more precise KMC and CI estimates from studies with higher sample sizes.

APPENDIX A

PROGRAM CODE

The program code in Appendices A.1 and A.2 were written and run in R 2.13.1.[11] The program code in Appendix A.1 first generates the survival data. Then it calculates the CI estimates, KMC estimates, and censoring proportions. Finally, it stores the estimates and censoring proportions in a text file for later use.

The program code in Appendix A.2 first retrieves the estimates from the files generated by the program code in Appendix A.1. Then it calculates the RD's. Finally, it outputs the KMC, CI, and RD at the median, 2.5%, and 97.5%. Optionally, the RD's can be graphed from years 1–10.

The sample code in Appendices A.1 and A.2 is used to test the parameter combination of $\lambda_1 = 0.015$, $\lambda_2 = 0.015$, $\lambda_C = 0.04$, and $N = 250$.

A.1 CUMULATIVE EVENT PROBABILITY ESTIMATION AND STORAGE

```
library(survival)
library(cmprsk)

setwd("C:\\Users\\Jeff\\Desktop\\Thesis\\Simulation Results")
getwd()

# Constants
nsim<-5000                                              # Number of models
est.times<-1:10                                         # Estimation times (Years 1-10)
ntimes<-length(est.times)

# Output matrix
output<-matrix(NA,nrow=nsim,ncol=2*ntimes+1)       # Cols. 1-10: CI; Cols. 11-20: KMC; Col. 21: Censor props.

# Model parameters
#############################################################################################
haz1<-0.015                                             # Primary event hazard
haz2<-0.015                                             # Competing event hazard
cens<-0.04                                              # Random censoring hazard
n<-250                                                  # Overall sample size
#############################################################################################

for(j in 1:nsim){                                       # Begin simulation loop

        ftime<-rep(NA,n)                                # Time vector
        event<-rep(NA,n)                                # Event vector

        ftime<-rexp(n,haz1+haz2)                        # Generate dependent exponential event times
        for(i in 1:n){                                  # Generate event types
                rand<-runif(1)
                event[i]<-ifelse(rand<=haz1/(haz1+haz2),1,2)
        }
        ctime<-rexp(n,cens)                             # Generate independent exponential censoring times
        event<-ifelse(ctime<ftime,0,event)
        Time<-apply(cbind(ftime,ctime),1,min)

        if(!is.na(table(event)[3])){                    # If at least 1 primary event
                fit.cuminc<-cuminc(Time,event)          # Cumulative incidence model
```

```
                fit.coxph<-coxph(Surv(Time,(event==1))~1)  # Cox PH model for primary event

                cuminc.all<-timepoints(fit.cuminc,est.times)
                cuminc.est<-cuminc.all$est["1 1",]         # Estimate CI for primary event

                info.coxph<-summary(survfit(fit.coxph),time=est.times)
                coxph.est<-1-info.coxph$surv               # Estimate KMC for primary event

                # Append max. estimates to CI and KMC if value is NA
                cuminc.max<-timepoints(fit.cuminc,max(Time[event==1]))$est["1 1",]
                cuminc.est[which(is.na(cuminc.est))]<-cuminc.max
                coxph.max<-1-summary(survfit(fit.coxph),time=max(Time[event==1]))$surv
                coxph.est<-c(coxph.est,rep(coxph.max,ntimes-length(coxph.est)))
        } else{                                            # Otherwise all KMC and CI estimates are 0
                cuminc.est<-rep(0,ntimes)
                coxph.est<-rep(0,ntimes)
        }

        output[j,1:ntimes]<-cuminc.est                     # Store CI estimates
        output[j,(1+ntimes):(2*ntimes)]<-coxph.est         # Store KMC estimates
        output[j,(2*ntimes+1)]<-table(event)[1]/n          # Store censoring proportions
}                                                          # End simulation loop

# Write output matrix to file
write(t(output),file=sprintf('haz1=%s, haz2=%s, cens=%s, n=%s.txt',haz1,haz2,cens,n),ncol=2*ntimes+1)
```

A.2 RELATIVE DIFFERENCE CALCULATION AND GRAPHING

```
library(survival)
library(cmprsk)

setwd("C:\\Users\\Jeff\\Desktop\\Thesis\\Simulation Results")
getwd()

# Constants
nsim<-5000                                              # Number of models
est.times<-1:10                                         # Estimation times (Years 1-10)
ntimes<-length(est.times)

rd<-matrix(NA,nrow=nsim,ncol=ntimes)                    # Relative difference matrix

ci.low<-rep(NA,ntimes)                                  # CI, KMC, and relative diff. vectors
ci.med<-rep(NA,ntimes)
ci.hi<-rep(NA,ntimes)
kmc.low<-rep(NA,ntimes)
kmc.med<-rep(NA,ntimes)
kmc.hi<-rep(NA,ntimes)
rd.low<-rep(NA,ntimes)
rd.med<-rep(NA,ntimes)
rd.hi<-rep(NA,ntimes)

# Model parameters
#############################################################################################
haz1<-0.015                                             # Primary event hazard
haz2<-0.015                                             # Competing event hazard
cens<-0.04                                              # Random censoring hazard
n<-250                                                  # Overall sample size
#############################################################################################

# Read output from file
output<-read.table(file=sprintf('haz1=%s, haz2=%s, cens=%s, n=%s.txt',haz1,haz2,cens,n))

for(i in 1:nsim){                                       # Calculate relative differences
        for(j in 1:ntimes){
                rd[i,j]<-(output[i,j+ntimes]-output[i,j])/output[i,j]
                if(is.nan(rd[i,j])) rd[i,j]<-0          # If CI = 0, RD = 0
```

```
        }
}

for(i in 1:ntimes){                                          # Low, medium, and high estimates
        ci.low[i]<-quantile(output[,i],0.025)
        ci.med[i]<-quantile(output[,i],0.5)
        ci.hi[i]<-quantile(output[,i],0.975)
        kmc.low[i]<-quantile(output[,i+ntimes],0.025)
        kmc.med[i]<-quantile(output[,i+ntimes],0.5)
        kmc.hi[i]<-quantile(output[,i+ntimes],0.975)
        rd.low[i]<-quantile(rd[,i],0.025)
        rd.med[i]<-quantile(rd[,i],0.5)
        rd.hi[i]<-quantile(rd[,i],0.975)
}
censor.low<-quantile(output[,2*ntimes+1],0.025)
censor.med<-quantile(output[,2*ntimes+1],0.5)
censor.hi<-quantile(output[,2*ntimes+1],0.975)


# Output CI, KMC, and RD for year 2: Median (2.5%-97.5%)
sprintf('%s (%s-%s)',round(ci.med[.2*ntimes],3),round(ci.low[.2*ntimes],3),round(ci.hi[.2*ntimes],3))
sprintf('%s (%s-%s)',round(kmc.med[.2*ntimes],3),round(kmc.low[.2*ntimes],3),round(kmc.hi[.2*ntimes],3))
sprintf('%s (%s-%s)',round(rd.med[.2*ntimes],3),round(rd.low[.2*ntimes],3),round(rd.hi[.2*ntimes],3))

# Output CI, KMC, and RD for year 5: Median (2.5%-97.5%)
sprintf('%s (%s-%s)',round(ci.med[.5*ntimes],3),round(ci.low[.5*ntimes],3),round(ci.hi[.5*ntimes],3))
sprintf('%s (%s-%s)',round(kmc.med[.5*ntimes],3),round(kmc.low[.5*ntimes],3),round(kmc.hi[.5*ntimes],3))
sprintf('%s (%s-%s)',round(rd.med[.5*ntimes],3),round(rd.low[.5*ntimes],3),round(rd.hi[.5*ntimes],3))

# Output CI, KMC, and RD for year 10: Median (2.5%-97.5%)
sprintf('%s (%s-%s)',round(ci.med[ntimes],3),round(ci.low[ntimes],3),round(ci.hi[ntimes],3))
sprintf('%s (%s-%s)',round(kmc.med[ntimes],3),round(kmc.low[ntimes],3),round(kmc.hi[ntimes],3))
sprintf('%s (%s-%s)',round(rd.med[ntimes],3),round(rd.low[ntimes],3),round(rd.hi[ntimes],3))

# Censoring proportion: Median (2.5%-97.5%)
sprintf('%s (%s-%s)',censor.med,censor.low,censor.hi)

##############################################################################################

# Graph relative differences
plot(est.times,rd.med,type="b",xlim=c(0,10),ylim=c(0,0.15),xlab="Years",
```

```
ylab="Relative difference (RD) between KMC and CI estimates")
title(paste("Relative Differences:
Primary event hazard =",haz1,"and Competing event hazard =",haz2,"
Censoring hazard =",cens,"and Sample size =",n))
lines(est.times,rd.low)
lines(est.times,rd.hi)
```

BIBLIOGRAPHY

1. Beyersmann, J. et al. (2009). "Simulating competing risks data in survival analysis." Statistics in Medicine **28**: 956-971.

2. Dignam, J. J., Kocherginsky, M. N. (2012). "Choice and Interpretation of Statistical Tests Used When Competing Risks Are Present." Journal of Clinical Oncology **26**(24): 4027-4034.

3. Fine, J. P., Gray, R. J. (1999). "A Proportional Hazards Model for the Subdistribution of a Competing Risk." Journal of the American Statistical Association **94**(446): 496-509.

4. Friedlin, B., Korn, E. L. (2004). "Testing treatment effects in the presence of competing risks." Statistics in Medicine **24**: 1703-1712.

5. Gaynor, J. J. et al. (1993). "On the Use of Cause-Specific Failure and Conditional Failure Probabilities: Examples from Clinical Oncology Data." Journal of the American Statistical Association **88**(422): 400-409.

6. Gooley, T. A. et al. (1999). "Estimation of Failure Probabilities in the Presence of Competing Risks: New Representations of Old Estimators." Statistics in Medicine **18**: 695-706.

7. Gray, R. J. (1988). "A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk." The Annals of Statistics **16**(3): 1141-1154.

8. Kay, Richard. (1986). "Treatment Effects in Competing-Risks Analysis of Prostate Cancer Data." Biometrics **42**: 203-211.

9. Klein, J. P., Moeschberger, M. L. (2003). Survival Analysis: Techniques for Censored and Truncated Data. 2nd ed. New York, Springer.

10. Mamounas, E. P., Anderson, S. J., et al. (2012). "Predictors of Locoregional Recurrence After Neoadjuvant Chemotherapy: Results From Combined Analysis of National Surgical Adjuvant Breast and Bowel Project B-18 and B-27." Journal of Clinical Oncology **30**(32): 3960-3966.

11. R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.