Published in final edited form as:

Stat Med. 2015 January 30; 34(2): 265–280. doi:10.1002/sim.6353.

Assessing treatment benefit with competing risks not affected by the randomized treatment

Edward L. Korn^{a,*}, James J. Dignam^a, and Boris Freidlin^a

^aBiometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD 20892, USA

^bDepartment of Public Health Sciences, University of Chicago, Chicago, IL 60637, USA, and NRG Oncology Statistics and Data Management, Philadelphia, PA 19103, USA

Abstract

The comparison of overall survival curves between treatment arms will always be of interest in a randomized clinical trial involving a life-shortening disease. In some settings, the experimental treatment is only expected to affect the deaths caused by the disease, and the proportion of deaths caused by the disease is relatively low. In these settings, the ability to assess treatment-effect differences between Kaplan-Meier survival curves can be hampered by the large proportion of deaths in both arms that are unrelated to the disease. To address this problem, frequently displayed are cause-specific survival curves or cumulative incidence curves, which respectively censor and immortalize events (deaths) not caused by the disease. However, the differences between the experimental and control treatment arms for these curves overestimate the difference between the overall survival curves for the treatment arms, and thus could result in overestimation of the benefit of the experimental treatment for the patients. To address this issue, we propose new estimators of overall survival for the treatment arms that are appropriate when the treatment does not affect the non-disease related deaths. These new estimators give a more precise estimate of the treatment benefit, potentially enabling future patients to make a more informed decision concerning treatment choice. We also consider the case where an exponential assumption allows the simple presentation of mortality rates as the outcome measures. Applications are given for estimating overall survival in a prostate-cancer treatment randomized clinical trial, and for estimating the overall mortality rates in a prostate-cancer screening trial.

Keywords

cause-specific survival; competing risks; cumulative incidence curves; death rates; randomized clinical trials; survival curves

1. Introduction

With a life-shortening disease, the aim of a new experimental treatment is to improve how long patients live or their quality of life. Therefore, overall survival (time-to-death from any

*Correspondence to: Edward L. Korn, Biometric Research Branch MSC9735, National Cancer Institute, 9609 Medical Center Drive, Bethesda MD 20892, USA korne@ctep.nci.nih.gov.

cause) is an important endpoint for randomized clinical trials involving life-shortening diseases. In some settings, however, there may be many deaths on the trial that are not associated with the disease (or treatments). This can be because the population under study is relatively old, and/or deaths from the disease are relatively uncommon. These different causes of deaths are known as competing risks, and there are various approaches to analysis [1]. Overall survival remains an important endpoint in the competing risks setting, but the ability to see any differences in overall survival due to the randomized treatment assignment can be difficult because of the high proportion of deaths unrelated to the disease under study.

In the competing risks setting, there is the possibility of attempting to isolate the effect of the treatments on survival by using cause-specific survival or the cumulative incidence of disease-related deaths. Cause-specific survival treats non-disease related deaths as censored observations in the analyses, whereas cumulative incidence treats them as living forever. Formally, the hazard of death is considered the sum of two hazards, $\lambda_1(t) + \lambda_2(t)$, where the hazard of dying from the disease (cause 1) is $\lambda_1(t)$, and the hazard of dying from other reasons (cause 2) is $\lambda_2(t)$. The overall survival can be expressed as

$$S(t) = \exp \left[- \int_0^t [\lambda_1(u) + \lambda_2(u)] du \right]$$

The cause-specific survival for cause i is defined by

$$S_i(t) = \exp \left[- \int_0^t \lambda_i(u) du \right]$$

with $S_1(t)$ being of primary interest in the present situation. Cause-specific survival is best not thought of as a survival function [2], but instead as a useful transformation of the cumulative cause-specific hazard. The cumulative incidence for cause i can be expressed as

$$I_i(t) = \int_0^t \lambda_i(u) S(u) du$$

with $I_1(t)$ being of primary interest in the present situation. In some applications, the time-to-event distributions are assumed to be exponential. In this case, the cause-specific hazards $\lambda_1(t)$ and $\lambda_2(t)$ do not depend on time. The treatment effect is then often summarized using the cause-specific mortality rate λ_1 and the overall mortality rate $\lambda_1 + \lambda_2$ for the experimental and control treatments.

There is no canonical approach for modeling that a new treatment does not affect cause 2 deaths. The model we will use in this paper is that the experimental treatment does not affect the cause-specific hazard for cause 2; see Appendix A for further discussion of defining what is meant by a treatment not affecting cause 2 deaths. Using E and C superscripts to represent the experimental treatment and the control treatment, this assumption is $\lambda_2^E(t) = \lambda_2^C(t)$. In the next section, we will use this assumption to examine the relationship between treatment-arm differences in overall survival curves (or overall mortality rates),

which we will argue is of primary interest, and the corresponding differences in cause-specific survival curves and incidence curves. In section 3, we propose a new estimator of the overall survival curves for the control and experimental treatments when it is reasonable to assume that $\lambda_2^E(t) \equiv \lambda_2^C(t)$. When would we expect $\lambda_2^E(t) \equiv \lambda_2^C(t)$? Consider a randomized clinical trial comparing two sizes of excision margin (1cm vs 3cm) for early stage melanoma; we would not expect the treatments to differ in non-melanoma deaths. On the other hand, consider a randomized trial of a long-term chemotherapy regimen for breast cancer with known toxicity; it is easy to imagine that the therapy could affect non-breast cancer deaths. A detailed application of the proposed estimator is given in section 3 to a randomized trial of short-term androgen deprivation for patients with early prostate cancer, where the treatment would not be expected to affect deaths unrelated to prostate cancer (in part, because the treatment is short term). In section 4, we proposed an analogous estimator for the overall mortality rate when it is reasonable to assume $\lambda_2^E = \lambda_2^C$ with exponential distributions. An application is given to a randomized trial of screening for prostate cancer, where the screening would not be expected to affect deaths unrelated to prostate cancer (because deaths due directly or indirectly to the screening are considered prostate-cancer-related deaths). Section 5 presents some limited results concerning the bias of the proposed overall survival estimator under moderate departures from the assumption that the treatments do not differentially affect other causes of death. We end with a discussion of applications of this approach to estimating other time-to-event distributions and with a further discussion of the assumptions involved.

2. Treatment-arm differences between survival curves with competing risks

Overall survival curves, cause-specific survival curves, and cumulative incidence curves are estimating different things. Under the assumption that the treatments do not affect cause 2 deaths ($\lambda_2^E(t) \equiv \lambda_2^C(t)$), what can we say about the difference in cause-specific survival curves and cumulative incidence curves in relation to the difference in overall survival curves? The difference in cause-specific survival curves will overestimate the overall survival difference. This is because the cause-specific survival curves are treating patients who die from cause 2 as being alive (at the time of their censoring due to death) and therefore having potential benefit from the new treatment (which they cannot). Formally, this can be seen by the relationship between the curve differences:

$$S_1^E(t) - S_1^C(t) = \frac{1}{S_2(t)} [S^E(t) - S^C(t)]$$

which follows because $S^E(t) = S_1^E(t) S_2(t)$ and $S^C(t) = S_1^C(t) S_2(t)$.

The difference in cumulative incidence curves between the treatment curves will also overestimate the overall survival difference. This is because the incidence curves do not take into account the fact that since more patients are alive in the experimental arm (due to the reduction of cause 1 deaths), more patients are at risk of dying and will die from cause 2 in the experimental arm than in the control arm. Since these excess cause 2 deaths in the

experimental arm are not counted (as no cause 2 deaths are counted in the incidence analysis), the difference in incidence curves will overestimate the survival difference. Formally, this can be seen by using the following Lemma (which incidentally does not depend on the assumption that the treatment does not affect cause 2 deaths).

Lemma

$$I_1^C(t) - I_1^E(t) = [S^E(t) - S^C(t)] + [I_2^E(t) - I_2^C(t)]$$

Proof

See Appendix B.

With the assumptions that the treatment does not affect cause 2 deaths $[\lambda_2^E(t) \equiv \lambda_2^C(t)]$ and that the treatment is working (i.e., $S^E(u) - S^C(u) > 0$ for all u), then $I_2^E(t) - I_2^C(t) \geq 0$ and then it follows that $I_1^C(t) - I_1^E(t) \geq S^E(t) - S^C(t)$.

To get an idea of the magnitude of the overestimation of the OS difference by the cumulative incidence difference and the cause-specific survival difference, we consider an example: Consider the case of constant cause-specific hazards with $\lambda_1^E = -\log(0.9)/5$ and $\lambda_2^C = -\log(0.6)/5$, representing five-year cause-specific survival of 90% and 60% for the experimental and control treatments. Table 1 displays the OS difference and incidence difference as the five-year cause-specific survival for cause 2 ranges from 100% to 10%. When there are no cause 2 deaths (top line of Table 1), the OS difference and incidence difference are the same. With an increasing proportion of cause 2 deaths, the incidence difference is larger in both an absolute and relative sense (until practically everyone is dying from cause 2 deaths). This is generally true except for very late times where the cause-specific survival difference goes to zero but the difference in cumulative incidence curves does not (see Appendix C for a comparison of the overestimation of OS differences by the cumulative incidence difference versus the cause-specific difference.). The difference in cause-1 cause-specific survivals is 30% for all lines of Table 1.

Taking the extreme case of the last line of Table 1 demonstrates why we believe that the overall survival difference is a better estimate of the treatment benefit of the experimental treatment than difference in cumulative incidence or cause-specific survival: At five years, the difference in cumulative incidence and cause-specific survival corresponds to an improvement for the experimental treatment of 13.1% and 30%, respectively. A reader of a medical journal, without a comprehensive understanding of competing risks methods, might reasonably assume that these are measures of the proportion of patients who are being helped by the experimental treatment. However, the true proportion of patients that are helped by the new therapy at five years is 3%, since only 9% and 6% of the patients are alive at that time. Therefore, although mathematically correct, we find the cumulative incidence and cause-specific survival differences misleading as a measure of treatment benefit.

It is interesting to note that if all the cause-specific hazards are small, both the OS difference and cumulative incidence difference are approximately $(\lambda_1^C - \lambda_1^E) t$. For example, if the five-year cause-specific survivals are 99% and 96% for experimental and control treatments, then the relative difference between the OS difference and incidence difference is less than 5% as the five-year cause-specific survival for cause 2 ranges from 100% to 91%.

The orderings described between the treatment-arm differences in survival curves, cause-specific survival curves, and cumulative incidence curves apply only when $\lambda_2^E(t) \equiv \lambda_2^C(t)$. In addition, in any given data analysis, the orderings of the observed curves may not hold even when $\lambda_2^E(t) \equiv \lambda_2^C(t)$ because by random variation the observed cause-specific hazards for cause 2 deaths will generally not be the same in the treatment arms.

With an exponential distribution assumption for the time-to-events, mortality rates (deaths per person-years) are usually reported to display treatment effects. Under the assumption that the treatments do not affect cause 2 deaths, that the difference in overall mortality rates is equal to the difference in cause-specific mortality rates, viz.,

$$(\lambda_1^C + \lambda_2) - (\lambda_1^E + \lambda_2) = \lambda_1^C - \lambda_1^E$$

Thus, the difference in cause-specific mortality rates is a good representation of the difference in overall mortality rates. Another measure of treatment effect frequently used is the proportional reduction in mortality rates. For this measure, however, the proportional reduction in cause-specific mortality rates overestimates the proportional reduction in overall mortality rates:

$$\frac{\lambda_1^C - \lambda_1^E}{\lambda_1^C} = \left[\frac{\lambda_1^C + \lambda_2}{\lambda_1^C} \right] \frac{(\lambda_1^C + \lambda_2) - (\lambda_1^E + \lambda_2)}{\lambda_1^C + \lambda_2}$$

3. Proposed estimators of overall survival curves

Let $t_1 \ t_2 \ \dots \ t_n$ be the ordered on-study times for the control and experimental groups combined. Let $\delta_i, i=1, \dots, n$ be an event-status indicator, equaling 1 (2) if a cause 1 (2) event occurs at time t_i , and equaling 0 if the observation is censored at t_i . We assume an independent censoring mechanism (Reference 3, pp 52-54). Let $r_i, i=1, \dots, n$ be equal to C or E depending on which treatment group the observation is in. With this notation, the Kaplan-Meier (product-limit) estimators of the control and experimental treatment (overall) survival curves are given by:

$$\begin{aligned} \hat{S}_{km}^C(t) &= \prod_{t_i \leq t} \frac{n_i^C - I(\delta_i > 0, r_i = C)}{n_i^C} \\ \hat{S}_{km}^E(t) &= \prod_{t_i \leq t} \frac{n_i^E - I(\delta_i > 0, r_i = E)}{n_i^E} \end{aligned}$$

where n_i^C (n_i^E) is the number of observations in the control (experimental) group at risk at time t_i . The proposed estimators are given by

$$\hat{S}^C(t) = \prod_{t_i \leq t} \left[\frac{n_i^C - I(\delta_i=1, r_i=C)}{n_i^C} \right] \left[\frac{(n_i^C + n_i^E) - I(\delta_i=2)}{(n_i^C + n_i^E)} \right]$$

$$\hat{S}^E(t) = \prod_{t_i \leq t} \left[\frac{n_i^E - I(\delta_i=1, r_i=E)}{n_i^E} \right] \left[\frac{(n_i^C + n_i^E) - I(\delta_i=2)}{(n_i^C + n_i^E)} \right]$$

(Throughout, we treat tied event times as if the cause 1 events happened first, followed by the cause 2 events, followed by censored observations.) Note that $\hat{S}^C(t) = \hat{S}_1^C(t) \hat{S}_2(t)$ and $\hat{S}^E(t) = \hat{S}_1^E(t) \hat{S}_2(t)$, where $\hat{S}_1^C(t)$ and $\hat{S}_1^E(t)$ are the product-limit estimator of the cause-specific survival function for cause 1 for the control group and experimental group, and $\hat{S}_2(t)$ is the product-limit estimator of the cause-specific survival for cause 2 in the pooled sample. Therefore, standard computer software can be used to calculate the proposed estimators. These relationships also show that the proposed estimators are consistent for $S^C(t)$ and $S^E(t)$ when $\lambda_2^E(t) \equiv \lambda_2^C(t)$, since their components are consistent for $S_1^C(t)$, $S_1^E(t)$ and $S_2(t)$ and $S^C(t) = S_1^C(t) S_2(t)$ and $S^E(t) = S_1^E(t) S_2(t)$ when $\lambda_2^E(t) \equiv \lambda_2^C(t)$.

Since the proposed overall survival curves pool the cause 2 survival information, they have less variability than the Kaplan-Meier estimators. Most relevant to the discussion here, the difference between the experimental and control survival curves would be expected to be much less variable for the proposed curves than Kaplan-Meier curves. To arrive at a simple formula for this efficiency, we consider the asymptotic variances of the curve differences under the simple situation of constant cause-specific hazards, the same censoring distribution in the two treatment arms, and under the null hypothesis of no difference between the treatment groups.

Proposition

Let $\lambda_1^E(t) \equiv \lambda_1^C(t) \equiv \lambda_1$, $\lambda_2^E(t) \equiv \lambda_2^C(t) \equiv \lambda_2$. Then

$$\frac{a.var \left[\hat{S}_{km}^E(t_0) - \hat{S}_{km}^C(t_0) \right]}{a.var \left[\hat{S}^E(t_0) - \hat{S}^C(t_0) \right]} = \frac{\log S_1(t_0) + \log S_2(t_0)}{\log S_1(t_0)}$$

Proof

See Appendix D.

As an example, consider a situation where the cause-specific survivals for the cause of interest (at a given time point) are approximately 0.95, and the cause-specific survivals for other causes are approximately 0.75. The proposition suggests that the variability of the difference in Kaplan-Meier curves at this time point will be approximately 6.6 times larger than the difference in the proposed overall survival curves. On the other hand, if the cause-specific survival were similar for the cause of interest and other causes, one would expect an

approximately doubling in the variance of the survival curves difference when using Kaplan-Meier curves instead of the proposed curves.

Greenwood-formula type variance estimators for the proposed survival curves are given by

$$\begin{aligned} \hat{v}ar \hat{S}^E &= \left[\hat{S}^E(t) \right]^2 \left[\sum_{t_i \leq t} \frac{I(\delta_i=1, r_i=E)}{n_i^E(n_i^E-1)} + \sum_{t_i \leq t} \frac{I(\delta_i=2)}{(n_i^C+n_i^E)(n_i^C+n_i^E-1)} \right] \\ \hat{v}ar \hat{S}^C &= \left[\hat{S}^C(t) \right]^2 \left[\sum_{t_i \leq t} \frac{I(\delta_i=1, r_i=C)}{n_i^C(n_i^C-1)} + \sum_{t_i \leq t} \frac{I(\delta_i=2)}{(n_i^C+n_i^E)(n_i^C+n_i^E-1)} \right] \end{aligned}$$

These variance estimators can be expressed in terms of the Greenwood-formula variance estimators $\hat{v}ar \hat{S}_1^C(t)$, $\hat{v}ar \hat{S}_2^C(t)$ and $\hat{v}ar \hat{S}_2(t)$ (which can be obtained with standard computer software)

$$\begin{aligned} \hat{v}ar \hat{S}^E(t) &= \left[\hat{S}_2(t) \right]^2 \hat{v}ar \hat{S}_1^E(t) + \left[\hat{S}_1^E(t) \right]^2 \hat{v}ar \hat{S}_2(t) \\ \hat{v}ar \hat{S}^C(t) &= \left[\hat{S}_2(t) \right]^2 \hat{v}ar \hat{S}_1^C(t) + \left[\hat{S}_1^C(t) \right]^2 \hat{v}ar \hat{S}_2(t) \end{aligned}$$

Since $\hat{S}^C(t)$ and $\hat{S}^E(t)$ are correlated (unlike $\hat{S}_{km}^C(t)$ and $\hat{S}_{km}^E(t)$), an analytic variance estimator for the difference between the proposed curves (at a given timepoint) is complex; we recommend using a bootstrap if an estimator of this variance is needed.

Noting that the null hypothesis $S^E(t) \equiv S^C(t)$ is equivalent to the null hypothesis $S_1^E(t) \equiv S_1^C(t)$ when $\lambda_2^E(t) \equiv \lambda_2^C(t)$, one can test $S^E(t) \equiv S^C(t)$ by using a standard test (e.g., log-rank test) on the cause-specific survival curves for cause 1. The p-value from this cause-specific log-rank test is the one that should be displayed with the proposed overall survival curves, and has been previously recommended as an appropriate test to use in this competing risks framework [4].

The assumption that $\lambda_2^E(t) \equiv \lambda_2^C(t)$ can be assessed by examining the observed cause-specific hazards $\lambda_2^E(t)$ and $\lambda_2^C(t)$, the cumulative hazards for cause 2, or $S_2^E(t)$ and $S_2^C(t)$ to see how different they are. However, these assessments can only identify large violations of the assumption, so that it is still necessary to have strong clinical/biological rationale that treatment does not affect cause 2 events to use the proposed overall survival estimators.

Application

RTOG-9408 conducted by the Radiation Therapy Oncology Group (RTOG) was an RCT of radiotherapy therapy versus radiotherapy plus short-term androgen-deprivation therapy (ADT) for early prostate cancer; OS was the primary endpoint [5]. The addition of ADT (for its four-month course post-radiation) would be expected to affect only prostate-cancer-specific deaths. The Kaplan-Meier overall survival curves are displayed in Figure 1A, demonstrating significant benefit of the addition of the ADT ($p=0.031$). Figure 1B displays the proposed estimates of the overall survival curves ($p=0.0006$). This p-value, which is the

same as the p-value for testing the difference in the cause-specific survival curves (Figure 2A), is smaller than the p-value associated with the Kaplan-Meier curves even though the proposed curves are closer together. Note the reduced variability in the difference between the proposed curves as compared to the Kaplan-Meier curves: the 5, 8, and 12 year estimates for the difference in the curves are $3.36\% \pm 1.64\%$, $5.47\% \pm 2.18\%$, and $5.26\% \pm 3.18\%$ for the Kaplan-Meier curves, and $1.00\% \pm 0.56\%$, $2.70\% \pm 0.82\%$, and $4.09\% \pm 1.55\%$ for the proposed curves (Table 2). The large standard errors for the Kaplan-Meier curves means that reliably quantifying the treatment benefit from these curves is problematic, whereas the proposed curves have much smaller standard errors. Figure 2B, which displays the cause-specific curves for non-prostate cancer deaths, shows there is no gross violation of the assumption $\lambda_2^E(t) \equiv \lambda_2^C(t)$. Assuming proportional hazards, the estimated hazard ratio is 0.93 with 95% confidence interval (0.80, 1.09).

For comparison, we also consider the cumulative incidence of prostate cancer deaths in the bottom of Table 1. The differences are similar but slightly larger than the differences in overall survival as estimated by the proposed method.

4 Proposed estimators of overall death rates (using exponential assumption)

The exponential assumption allows a simple presentation of death rates as outcome measures in the treatment groups; it is routinely used in certain clinical settings with rare events such as screening. Using the notation of the previous section, the log-likelihood can be expressed in terms of the cause-specific (constant) mortality rates:

$$\begin{aligned} \log L &= d_1^C \log \lambda_1^C + d_2^C \log \lambda_2 - \left(\lambda_1^C + \lambda_2 \right) \sum_{i=1}^n t_i (r_i = C) \\ &\quad + d_1^E \log \lambda_1^E + d_2^E \log \lambda_2 - \left(\lambda_1^E + \lambda_2 \right) \sum_{i=1}^n t_i (r_i = E) \end{aligned}$$

where d_1^C , d_2^C , d_1^E , and d_2^E are the number of cause 1 and cause 2 events (deaths) in the control and experimental groups, e.g., $d_1^C = \sum_{i=1}^n I(\delta_i = 1, r_i = C)$. The maximum likelihood estimators of the mortality rates and their variances are given by

$$\begin{aligned} \hat{\lambda}_1^C &= \frac{d_1^C}{\sum_{i=1}^n t_i I(r_i = C)} & \text{var}(\hat{\lambda}_1^C) &= \frac{(\hat{\lambda}_1^C)^2}{d_1^C} \\ \hat{\lambda}_1^E &= \frac{d_1^E}{\sum_{i=1}^n t_i I(r_i = E)} & \text{var}(\hat{\lambda}_1^E) &= \frac{(\hat{\lambda}_1^E)^2}{d_1^E} \\ \hat{\lambda}_2 &= \frac{d_2^C + d_2^E}{\sum_{i=1}^n t_i} & \text{var}(\hat{\lambda}_2) &= \frac{(\hat{\lambda}_2)^2}{d_2^C + d_2^E} \end{aligned}$$

with the mortality rate estimators being asymptotically uncorrelated. (These are the standard estimators for λ_1^C and λ_1^E and their variances [6].) The proposed (and maximum likelihood) estimators for the overall mortality rates, difference in mortality rates, proportional reduction in mortality rates, and relevant variance estimators are given by

$$\begin{aligned}\hat{\lambda}^C &= \hat{\lambda}_1^C + \hat{\lambda}_2 & \text{var}(\hat{\lambda}^C) &= \frac{(\hat{\lambda}_1^C)^2}{d_1^C} + \frac{(\hat{\lambda}_2)^2}{d_2^C + d_2^E} \\ \hat{\lambda}^E &= \hat{\lambda}_1^E + \hat{\lambda}_2 & \text{var}(\hat{\lambda}^E) &= \frac{(\hat{\lambda}_1^E)^2}{d_1^E} + \frac{(\hat{\lambda}_2)^2}{d_2^C + d_2^E} \\ \text{var}(\hat{\lambda}^C - \hat{\lambda}^E) &= \frac{(\hat{\lambda}_1^C)^2}{d_1^C} + \frac{(\hat{\lambda}_1^E)^2}{d_1^E} \\ \text{var}\left(\frac{\hat{\lambda}^C - \hat{\lambda}^E}{\hat{\lambda}^C}\right) &= \frac{(\hat{\lambda}_1^E + \hat{\lambda}_2)^2 (\hat{\lambda}_1^C)^2}{(\hat{\lambda}_1^C + \hat{\lambda}_2)^4} \frac{1}{d_1^C} + \frac{(\hat{\lambda}_1^E)^2}{(\hat{\lambda}_1^C + \hat{\lambda}_2)^2} \frac{1}{d_1^E} + \frac{(\hat{\lambda}_1^C - \hat{\lambda}_1^E)^2 (\hat{\lambda}_2)^2}{(\hat{\lambda}_1^C + \hat{\lambda}_2)^4} \frac{1}{d_2^C + d_2^E}\end{aligned}$$

These can be compared to the usual estimators of the overall mortality rates:

$$\begin{aligned}\hat{\lambda}_{usual}^C &= \frac{d_1^C + d_2^C}{\sum_{i=1}^n t_i I(r_i = C)} & \text{Var}(\hat{\lambda}_{usual}^C) &= \frac{(\hat{\lambda}_{usual}^C)^2}{d_1^C + d_2^C} \\ \hat{\lambda}_{usual}^E &= \frac{d_1^E + d_2^E}{\sum_{i=1}^n t_i I(r_i = E)} & \text{Var}(\hat{\lambda}_{usual}^E) &= \frac{(\hat{\lambda}_{usual}^E)^2}{d_1^E + d_2^E} \\ \text{var}(\hat{\lambda}_{usual}^C - \hat{\lambda}_{usual}^E) &= \frac{(\hat{\lambda}_{usual}^C)^2}{d_1^C + d_2^C} + \frac{(\hat{\lambda}_{usual}^E)^2}{d_1^E + d_2^E} \\ \text{var}\left(\frac{\hat{\lambda}_{usual}^C - \hat{\lambda}_{usual}^E}{\hat{\lambda}_{usual}^C}\right) &= \left(\frac{\hat{\lambda}_{usual}^E}{\hat{\lambda}_{usual}^C}\right)^2 \left(\frac{1}{d_1^E + d_2^E} + \frac{1}{d_1^C + d_2^C}\right)\end{aligned}$$

With $\lambda_1^E = \lambda_1^C (= \lambda_1)$, we have simple expressions for the ratio of the asymptotic variances of the usual and proposed estimators of the difference and proportional reduction in mortality rates:

$$\begin{aligned}\frac{a.\text{var}(\hat{\lambda}_{usual}^C - \hat{\lambda}_{usual}^E)}{a.\text{var}(\hat{\lambda}^C - \hat{\lambda}^E)} &= \frac{\lambda_1 + \lambda_2}{\lambda_1} \\ \frac{a.\text{var}\left[\left(\frac{\hat{\lambda}_{usual}^C - \hat{\lambda}_{usual}^E}{\hat{\lambda}_{usual}^C}\right) / \hat{\lambda}_{usual}^C\right]}{a.\text{var}\left[\left(\frac{\hat{\lambda}^C - \hat{\lambda}^E}{\hat{\lambda}^C}\right) / \hat{\lambda}^C\right]} &= \frac{\lambda_1 + \lambda_2}{\lambda_1}\end{aligned}$$

These asymptotic results follow from the consistency of the hazard rate estimators and the delta method. Note that the variance ratio is the same for these exponential-distribution estimators as it was for the product-limit estimators given in section 3.

Application

The European Randomized Study of Screening for Prostate Cancer [7] was a trial of PSA screening for prostate cancer and found a reduction in prostate-cancer mortality: 0.39 and

0.50 prostate-cancer deaths per 1000 person-years for the screening and control groups, respectively (Table 3). This corresponds to an absolute difference of -0.10 per 1000 person-years and a relative reduction of 21%. The overall mortality rates were 18.21 and 18.49 deaths per 1000 person-years for the screening and control groups, respectively, corresponding to an absolute difference of -0.284 ± 0.209 and a relative reduction of $1.53\% \pm 1.12\%$ (Table 3). Using our proposed estimators (which assume that the screening does not affect non-prostate-cancer mortality), the overall mortality rates are 18.31 and 18.41 deaths per 1000 person-years for the screening and control groups, respectively, corresponding to an absolute difference of -0.104 ± 0.032 and a relative reduction of $0.56\% \pm 0.17\%$. As noted above, this absolute difference is the same as the absolute difference in prostate-cancer death rates. (The relative reduction in prostate-cancer mortality (21%) is much larger than the relative reduction in overall mortality, 0.56% or 1.53% using the usual estimator). The standard errors for the differences of the usual mortality estimators are more than six times larger than the standard errors for the proposed estimators. As a result, the usual estimators and standard errors cannot exclude a possible detriment of the screening, while the proposed estimators can. As was done with this study, it is important in this type of application to treat deaths due directly or indirectly to the screening as disease-related deaths.

5. Robustness of proposed overall survival estimator

The proposed estimators were derived under the assumption that the cause 2 hazards are the same in both treatment arms, and, under that assumption the estimators are consistent for the true overall survival curves (section 3). We briefly consider the situation when the proposed estimators are used and the assumption is moderately not satisfied. To do this, we reconsider the scenario described in Table 1 (constant cause-specific hazards) and estimate the survival curve differences when the assumption is not satisfied.

With increasing sample size and π proportion of the patients randomized to the experimental arm, the proposed estimators converge as follows:

$$\begin{aligned}\hat{S}^C(t) &\rightarrow e^{-\lambda_1^C t} S_2(t) \\ \hat{S}^E(t) &\rightarrow e^{-\lambda_1^E t} S_2(t)\end{aligned}$$

where

$$S_2(t) = \exp \left[- \int_0^t \frac{\lambda_2^E \pi e^{-(\lambda_1^E + \lambda_2^E)u} + \lambda_2^C (1 - \pi) e^{-(\lambda_1^C + \lambda_2^C)u}}{\pi e^{-(\lambda_1^E + \lambda_2^E)u} + (1 - \pi) e^{-(\lambda_1^C + \lambda_2^C)u}} du \right]$$

Table 4 displays the difference in treatment-arm curves with differing amounts of model assumption violation (with $\pi=1/2$). As expected, the treatment-arm difference in the proposed overall survival estimators is biased for the true treatment-arm difference in overall survival curves when the cause-specific hazards for cause 2 are different between

treatment arms. Note, however, that the difference in cause-specific curves and cumulative incidence curves also do not estimate the true overall survival difference.

6. Discussion

In some applications, a time-to-event endpoint other than overall survival may be highly relevant to patients for comparing treatment arms, and the proposed methods can be used for these other endpoints. For example, in comparing adjuvant treatments for early breast cancer, patients who are ostensibly free of disease after initial treatment are randomized between the adjuvant treatment arms. In this situation, there may be very few breast cancer deaths even with long follow-up, and disease-free survival (time from randomization to recurrence of breast cancer or death) may be an appropriate outcome for patients to judge the clinical benefit of a new treatment (as it may measure direct quality of life benefit). If the treatments are expected to affect only the hazard of breast cancer recurrence (and breast cancer death) and not the hazard of dying from other causes, the estimators proposed in this paper for overall survival can easily be modified to estimate disease-free survival. Another potential application is in randomized screening trials where time to symptomatic or metastatic disease may be a relevant outcome for patients in addition to overall survival. For example, in the European Randomized Study of Screening for Prostate Cancer discussed in section 4, the affect of screening on time to metastatic prostate cancer was evaluated as an outcome [8]. If the screening is expected to affect only the development of metastatic prostate (and prostate cancer death) and not the hazard of dying from other causes, the estimators proposed in this paper can easily be modified to estimate metastases-free survival.

The use of the proposed survival estimators rests on the assumption that the treatments do not affect non-disease related deaths, whereas the usual Kaplan-Meier curves (or overall death rates) require no assumptions. If the assumption is not satisfied, the difference in the proposed survival estimators will be a biased estimator of the difference in the true overall survival curves (Table 4). Why use the proposed estimators? The main reason is that the difference in Kaplan-Meier curves (or the proportional reduction in mortality) can be so variable as to be almost meaningless when a high proportion of deaths are not related to the disease or treatments. Perhaps because of this, sometimes overall survival curves are not presented at all in trial reports. Instead, cause-specific survival curves or cumulative incidence curves are presented. However, the use of cause-specific survival curves or cumulative incidence curves as a measure of treatment benefit implicitly assumes that the treatments do not affect non-disease related deaths (as does the proposed method); if the treatments did affect other causes of death, then the difference in cause-specific survival curves or cumulative incidence curves would be misleading if interpreted as a measure of treatment benefit. In addition, although cause-specific curves can be useful for demonstrating there is some treatment effect [9], they have no easy interpretation. Cumulative incidence curves do have an easy interpretation, but they overestimate the treatment benefit for patients (even when the treatments do not affect non-disease related deaths), especially when the proportion of cause-specific deaths is small relative to the overall number of deaths.

With an exponential assumption, the difference in cause-specific rates does estimate the difference in overall mortality rates (provided the treatments do not affect the other causes of death). The proportional reduction in cause-specific rates, however, will overestimate the proportional reduction in overall rates. With or without an exponential assumption, another reason to prefer overall survival estimators over cause-specific or cumulative incidence estimators is that they display the treatment-derived survival benefit in the context of the overall survival prognosis of the patient population of interest. For example, it is possible that a 15% difference in three-year survival from 80% to 95% will have a different meaning to a patient than a 15% difference from 10% to 25%. Incidence curves do not contain that extra information. Because of these reasons, the proposed overall survival estimators, in addition to the usual Kaplan-Meier curves (or death rates), can be a useful presentation of the benefits of a treatment in the competing risks setting when the proportion of competing-risk events is large.

Acknowledgements

J. Dignam was supported by Public Health Service grants U10 CA21661 and U10 CA180822 from the National Cancer Institute, NIH, U.S. Department of Health and Human Services.

Appendix A

Defining what is meant by “the treatment does not affect cause 2 events” is not straightforward. We have defined this to mean that the treatment does not affect the cause-specific hazard for cause 2. This is not a completely satisfactory definition: Imagine a population that is composed to high risk and low risk subpopulations of patients. A treatment that affects the cause-1 hazard but not the cause-2 hazard in each of the subpopulations, will affect the cause-2 hazard in the overall population. This can be seen by noting that the cause-specific hazard for cause 2 at a specific time t_0 for the overall population is a mixture of the cause-2 specific hazards for the two subpopulations, with the mixing proportion depending on the probabilities a patient is in each of the subpopulations given that he is alive at time t_0 . These latter probabilities depend on the cause-specific hazards for both causes, so that if the cause-specific hazard for cause 1 is affected by the treatment in the subpopulations, the overall cause-specific hazard for cause 2 will be affected. Because of this difficulty, it would perhaps be better to state the assumption as “the treatment does not, on average, affect cause 2 events” to reinforce the notion that the definition is applying to a specific population.

To give an example of the potential bias due to heterogeneity, consider a hypothetical population consisting of one-half a high-risk subpopulation and one-half a low-risk subpopulation, for which the treatment does not affect the cause 2 deaths in either subpopulation: For the high risk subpopulation, $\lambda_1^E = -\log(0.9)/5$, $\lambda_1^C = -\log(0.6)/5$, and $\lambda_2^E = \lambda_2^C = -\log(0.5)/5$, and for the low-risk subpopulation, $\lambda_1^E = -\log(0.95)/5$, $\lambda_1^C = -\log(0.8)/5$, and $\lambda_2^E = \lambda_2^C = -\log(0.75)/5$. For the overall population, the five-year overall survival difference between the treatment arms is 13.13%, whereas using the proposed method (which incorrectly assumes that the cause-specific hazard for cause 2

deaths in the overall population does not depend on treatment arm) yields a treatment-arm difference of 13.72%, a minor difference.

An alternative approach to defining what is meant by a treatment only affecting deaths from cause 1 is through a latent failure time model [10]. Let T_1 and T_2 be the latent failure times associated with the two causes of death, with the observed data being $T = \min(T_1, T_2)$ and δ an indicator of the cause of death. A possible model for the treatment only affecting cause 1 is

$$S^C(t_1, t_2) = S^E(g(t_1), t_2)$$

where $S^C(t_1, t_2)$ and $S^E(t_1, t_2)$ are the bivariate survival distributions of the latent failure times for the control and experimental treatment, respectively, and $g(t)$ represents the experimental treatment being better than the control treatment. The overall survival curves for the control and experimental treatment are given by $S^C(t) = S^C(t, t)$ and $S^E(t) = S^E(t, t)$. The main problem with latent failure time models is the non-identifiability of the correlation of T_1 and T_2 from the survival data [11]. In particular, the assumption that these latent variables are independent cannot be assessed from the data. Even with an assumed latent-failure statistical model, inferences about cause-specific survival as representing survival with other causes of death removed are unsupportable [12]. For these reasons, we have focused on observable quantities in this paper.

Appendix B

Lemma

$$I_1^C(t) - I_1^E(t) = [S^E(t) - S^C(t)] + [I_2^E(t) - I_2^C(t)]$$

Proof

$$\begin{aligned} I_1^C(t) - I_1^E(t) &= \int_0^t \lambda_1^C(u) S^C(u) du - \int_0^t \lambda_1^E(u) S^E(u) du \\ &= \int_0^t [\lambda_1^C(u) + \lambda_2^C(u)] S^C(u) du - \int_0^t [\lambda_1^E(u) + \lambda_2^E(u)] S^E(u) du - \int_0^t \lambda_2^C(u) du + \int_0^t \lambda_2^E(u) S^E(u) du \\ &= [1 - S^C(t)] - [1 - S^E(t)] + [I_2^E(t) - I_2^C(t)] \end{aligned}$$

Appendix C

Since both the difference in incidence and the difference in cause-specific survival curves overestimate the OS difference, we investigate which has worse overestimation. The overestimation of the OS difference will tend to be larger for the cause-specific curves than the incidence curves when the difference in cause-specific curves $[S_1^E(u) - S_1^C(u)]$ is larger with larger times u :

Proposition

If $\frac{d}{du} [S_1^E(u) - S_1^C(u)] \geq 0$ for all $u \leq t$, then

$$S_1^E(t) - S_1^C(t) \geq I_1^C(t) - I_1^E(t)$$

Proof

$$\begin{aligned} S_1^E(t) - S_1^C(t) &= \int_0^t \frac{d}{du} [S_1^E(u) - S_1^C(u)] du = \int_0^t [\lambda_1^C(u) S_1^C(u) - \lambda_1^E(u) S_1^E(u)] du \\ &= \int_0^t \frac{1}{S_2(u)} \frac{d}{du} [I_1^C(u) - I_1^E(u)] du \geq \int_0^t \frac{d}{du} [I_1^C(u) - I_1^E(u)] du = I_1^C(t) - I_1^E(t) \end{aligned}$$

where the inequality follows because the integrands are non-negative.

Eventually, the survival curves go to 0 as t gets large, so eventually

$S_1^E(t) - S_1^C(t) \leq I_1^C(t) - I_1^E(t)$ for t large enough. The above proposition shows that this will not happen provided that distance between the cause-specific survival curves is still getting bigger. This may happen much later than when the distance between the overall survival curves starts becoming smaller. As an example, consider again the case of constant cause-specific hazards with $\lambda_1^E = -\log(0.9)/5$, $\lambda_1^C = -\log(0.6)/5$, and

$\lambda_2^E = \lambda_2^C = -\log(0.5)/5$ (line 6 of Table 1). With these parameters, the between-arm difference at five years in overall survival, cumulative incidence, and cause-specific survival

is 15%, 22.4%, and 30%, respectively. For this example, $\frac{d}{du} [S_1^E(u) - S_1^C(u)] \geq 0$ for

$u \leq 19.5$ years, but $\frac{d}{du} [S_1^E(u) - S_1^C(u)] < 0$ for $u > 19.5$ years. In fact, even when the cause-specific survival curves start coming back together (e.g., when the experimental treatment has delayed but not eliminated disease-related deaths), it may be a long time before

$S_1^E(t) - S_1^C(t) \leq I_1^C(t) - I_1^E(t)$; in this example, $S_1^E(u) - S_1^C(u) \geq I_1^C(u) - I_1^E(u)$ for $u \leq 58$ years. That is, for practical purposes, the difference in cause-specific curves is always larger than the difference in cumulative incidence curves for this example.

Appendix D

Proposition

Let $\lambda_1^E(t) \equiv \lambda_1^C(t) \equiv \lambda_1$, $\lambda_2^E(t) \equiv \lambda_2^C(t) \equiv \lambda_2$. Then

$$\frac{a.var [\hat{S}_{km}^E(t_0) - \hat{S}_{km}^C(t_0)]}{a.var [\hat{S}^E(t_0) - \hat{S}^C(t_0)]} = \frac{\log S_1(t_0) + \log S_2(t_0)}{\log S_1(t_0)}$$

Proof

Let $G(u)$ be the distribution function of the censoring distribution, assumed the same in the treatment arms. For a product-limit estimator $\hat{S}_Y(t)$ of a survival function based on data $\{\min(Y_i, C_i), \delta_i = I(Y_i < C_i), i=1, \dots, n\}$, the asymptotic variance of $\hat{S}_Y(t_0)$ is given by [13]

$$a.var \left[\hat{S}_Y(t_0) \right] = \frac{1}{n} [1 - F(t_0)]^2 \int_0^t \frac{f(u) d(u)}{[1 - F(u)]^2 [1 - H(u)]}$$

where $f(u)$ and $F(u)$ are the density and distribution function of Y , and $H(u)$ is the distribution function of C . Applying this result to the independent quantities $\hat{S}_{km}^E(t_0)$ and $\hat{S}_{km}^C(t_0)$, yields

$$a.var \left[\hat{S}_{km}^E(t_0) - \hat{S}_{km}^C(t_0) \right] = \left(\frac{1}{n^E} + \frac{1}{n^C} \right) [S(t_0)]^2 (\lambda_1 + \lambda_2) \int_0^{t_0} \frac{e^{(\lambda_1 + \lambda_2)u}}{1 - G(u)} du$$

(with the identifications $S_Y(u) \equiv S(u)$ and $H(u) \equiv G(u)$). Applying this result to the independent quantities $\hat{S}_1^E(t_0)$ and $\hat{S}_1^C(t_0)$, yields

$$a.var \left[\hat{S}_1^E(t_0) - \hat{S}_1^C(t_0) \right] = \left(\frac{1}{n^E} + \frac{1}{n^C} \right) [S_1(t_0)]^2 (\lambda_1) \int_0^{t_0} \frac{e^{(\lambda_1 + \lambda_2)u}}{1 - G(u)} du$$

(with the identifications $S_Y(u) \equiv S_1(u)$ and $H(u) \equiv 1 - [1 - G(u)]S_2(u)$). Because $\hat{S}_2(t_0)$ is asymptotically independent of $\hat{S}_1^E(t_0)$ and $\hat{S}_1^C(t_0)$ (see below), and

$$a.E \left[\hat{S}_1^E(t_0) - \hat{S}_1^C(t_0) \right] = 0,$$

$$\begin{aligned} a.var \left[\hat{S}_1^E(t_0) - \hat{S}_1^C(t_0) \right] &= a.var \left[\hat{S}_2(t_0) \left\{ \hat{S}_1^E(t_0) - \hat{S}_1^C(t_0) \right\} \right] \\ &= \left\{ a.E \left[\hat{S}_2(t_0) \right] \right\}^2 a.var \left[\hat{S}_1^E(t_0) - \hat{S}_1^C(t_0) \right] \\ &= [S_2(t_0)]^2 a.var \left[\hat{S}_1^E(t_0) - \hat{S}_1^C(t_0) \right] \\ &= \left(\frac{1}{n^E} + \frac{1}{n^C} \right) [S(t_0)]^2 (\lambda_1) \int_0^{t_0} \frac{e^{(\lambda_1 + \lambda_2)u}}{1 - G(u)} du \end{aligned}$$

because $S(t_0) = S_1(t_0)S_2(t_0)$. Therefore, the ratio of the asymptotic variances equals

$$\frac{\lambda_1 + \lambda_2}{\lambda_1} = \frac{\log S_1(t_0) + \log S_2(t_0)}{\log S_1(t_0)}$$

and this concludes the proof.

To demonstrate that $\hat{S}_1^E(t)$, $\hat{S}_1^C(t)$, and $\hat{S}_2(t)$ are asymptotically independent, consider the trivariate martingale $(M_{1i}^E(t), M_{1i}^C(t), M_{2i}(t))$ as applicable to each individual in the pooled sample ($i=1, \dots, n$) [14]:

$$\begin{aligned} M_{1i}^E(t) &= I(T_i \leq t, J_i=1) - \int_0^t \pi Y_i(u) \lambda_1(u) du \\ M_{1i}^C(t) &= I(T_i \leq t, J_i=2) - \int_0^t (1-\pi) Y_i(u) \lambda_1(u) du \\ M_{2i}(t) &= I(T_i \leq t, J_i=3) - \int_0^t Y_i(u) \lambda_2(u) du \end{aligned}$$

where π is the probability an individual is randomized to the experimental treatment arm, $J_i = 1, 2$, or 3 depending upon whether the i th individual has (1) a type 1 event and is in the experimental treatment arm, (2) a type 1 event and is in the control treatment arm, or (3) a type 2 event, respectively, $Y_i(u) = I(T_i \geq u)$ is assumed, and recall the null hypothesis

$S_1^E(t) \equiv S_1^C(t)$ is assumed. The sum of this trivariate martingale of the sample, $(M_{1\bullet}^E(t), M_{1\bullet}^C(t), M_{2\bullet}(t))$, is also a martingale, and because it is assumed that the survival times have densities and that the component martingales are orthogonal, the predictable variation process of $(M_{1\bullet}^E(t), M_{1\bullet}^C(t), M_{2\bullet}(t))$ has zeros on the off-diagonals [14]. The centered estimator of the estimated survival curves $(G_1^E(t), G_1^C(t), G_2(t))$ is also a trivariate martingale (Reference 3, p 170), where

$$\begin{aligned} G_1^E(t) &= \frac{\hat{S}_1^E(t)}{\hat{S}_1^E(0)} - 1 \\ G_1^C(t) &= \frac{\hat{S}_1^C(t)}{\hat{S}_1^C(0)} - 1 \\ G_2(t) &= \frac{\hat{S}_2(t)}{\hat{S}_2(0)} - 1 \end{aligned}$$

and will also predictable variation process with off-diagonal zeros [14]. By the multivariate martingale central limit theorem [14], $(G_1^E(t), G_1^C(t), G_2(t))$ converge to a trivariate normal distribution with zero correlations. Therefore, $\hat{S}_1^E(t)$, $\hat{S}_1^C(t)$ and $\hat{S}_2(t)$ are asymptotically independent.

References

1. Marubini, E.; Valsecchi, MG. Analysing Survival Data from Clinical Trials and Observational Studies. Wiley; Chichester: 1995. p. 331-363.
2. Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. Biometrics. 1978; 34:541–554. [PubMed: 373811]
3. Kalbfleisch, JD.; Prentice, RL. The Statistical Analysis of Failure Time Data. Second Edition.. Wiley; Hoboken, New Jersey: 2002.
4. Freidlin B, Korn EL. Testing treatment effects in the presence of competing risks. Statistics in Medicine. 2005; 24:1703–1712. [PubMed: 15706579]
5. Jones CU, Hunt D, McGowan DG, et al. Radiotherapy and short-term androgen deprivation for localized prostate cancer. New England Journal of Medicine. 2011; 365:107–118. [PubMed: 21751904]
6. Miller, RG. Survival Analysis. Wiley; New York: 1981. p. 24-25.

7. Schroder FH, Hugosson J, Roobol MJ, et al. Prostate-cancer mortality at 11 years of follow-up. *New England Journal of Medicine*. 2012; 366:981–990. [PubMed: 22417251]
8. Schroder FH, Hugosson J, Carlsson S, et al. Screening for prostate cancer decreases the risk of developing metastatic disease: findings from the European Randomized Study of Screening for Prostate Cancer. *European Urology*. 2012; 62:745–752. [PubMed: 22704366]
9. Freidlin B, Korn EL. Testing treatment effects in the presence of competing risks. *Statistics in Medicine*. 2005; 24:1703–1712. [PubMed: 15706579]
10. Korn EL, Dorey FJ. Applications of crude incidence curves. *Statistics in Medicine*. 1992; 11:813–829. [PubMed: 1594819]
11. Tsiatis A. A nonidentifiability aspect of the problem of competing risks. *Proc Nat Acad Sci*. 1975; 72:20–22. [PubMed: 1054494]
12. Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics*. 1978; 34:541–554. [PubMed: 373811]
13. Miller RG. What price Kaplan-Meier? *Biometrics*. 1983; 39:1077–1081. [PubMed: 6671119]
14. Andersen, PK.; Borgan, O.; Gill, RD.; Keiding, N. *Statistical Models Based on Counting Processes*. Springer-Verlag; New York: 1991. p. 72-83.

Figure 1. Overall survival curves for RTOG-9408 (solid lines=RAD+ADT, dotted lines=RAD). (A) Kaplan-Meier survival curves (B) Proposed survival curve.

Figure 2. Cause-specific survival curves for RTOG-9408 (solid lines=RAD+ADT, dotted lines=RAD). (A) For prostate-cancer deaths (B) For non-prostate cancer deaths

Table 1

Hypothetical overall survival difference and cumulative incidence difference between treatment arms at five years for exponential survival with five-year cause-specific survival of 90% for the experimental treatment and 60% for the control treatment

5-year cause-specific survival for other causes of death (^a)	Overall survival (OS) at 5 years			Cumulative incidence (for deaths of interest) at 5 years			Difference between OS difference and incidence difference	
	Experimental	Control	Difference	Control	Experimental	Difference	Absolute	Relative ^b (%)
100%	90%	60%	30%	40.0%	10.0%	30.0%	0.0%	0%
90%	81%	54%	27%	38.1%	9.5%	28.6%	1.6%	6%
80%	72%	48%	24%	36.2%	9.0%	27.2%	3.2%	13%
70%	63%	42%	21%	34.2%	9.4%	25.7%	4.7%	22%
60%	54%	36%	18%	32.0%	7.9%	24.1%	6.1%	34%
50%	45%	30%	15%	29.7%	7.3%	22.4%	7.4%	50%
40%	36%	24%	12%	27.2%	6.6%	20.6%	8.6%	72%
30%	27%	18%	9%	24.4%	5.9%	18.6%	9.6%	106%
20%	18%	12%	6%	21.2%	5.0%	16.2%	10.2%	169%
10%	9%	6%	3%	17.1%	4.0%	13.1%	10.1%	336%

^aThe cause-specific survival for other causes of death is the same for both the experimental and control treatment arms.

^b Absolute difference divided by overall survival difference.

Table 2

Overall survival results of RTOG-9408 (radiotherapy alone (RAD) versus RAD+ADT)

	Treatment Arm		Difference ^(a)
	RAD+ADT (n=987)	RAD (n=992)	
Deaths (% of randomized)	359 (36.4%)	404 (40.7%)	
Prostate cancer deaths	40 (4.1%)	74 (7.5%)	
Other deaths	319 (32.3%)	330 (33.3%)	
Kaplan-Meier overall survival (%±SE)			
5-years	85.97±1.13	82.62±1.22	3.36±1.64
8-years	72.78±1.49	67.32±1.56	5.47±2.18
12-years	51.15±2.27	45.88±2.29	5.27±3.18
Proposed overall survival (%±SE)			
5-years	84.79±0.85	83.78±0.90	1.00±0.56
8-years	71.37±1.11	68.67±1.19	2.70±0.82
12-years	50.54±1.74	46.46±1.85	4.08±1.55
Cumulative incidence of prostate cancer death (%±SE)			
5-years	1.07±0.34	2.10±0.42	-1.03±0.59
8-years	2.56±0.53	5.45±0.76	-2.89±0.94
12-years	6.13±1.10	11.16±1.45	-5.03±1.81

^aStandard errors for differences are from a bootstrap with 1000 bootstrap resamples.

Table 3Overall mortality results of the European Randomized Study of Screening for Prostate Cancer^(a)

	Treatment Arm		Difference	Proportional reduction
	Screening Group (n=72891)	Control group (n=89532)		
Deaths (% of randomized)	13917 (19.1%)	17256 (19.3%)		
Prostate cancer deaths	299 (0.4%)	462 (0.5%)		
Other deaths	13618 (18.7%)	16794 (18.8%)		
Person-years	764233	933053		
Prostate cancer death rate per 1000 person years ($\pm$ SE)	0.391 $\pm$ 0.023	0.495 $\pm$ 0.023	-0.104 $\pm$ 0.032	21.01% $\pm$ 0.05%
Overall death rate (usual calculation) per 1000 person years ($\pm$ SE)	18.210 $\pm$ 0.154	18.494 $\pm$ 0.141	-0.284 $\pm$ 0.209	1.53% $\pm$ 1.12%
Proposed overall death rate per 1000 person years ($\pm$ SE)	18.309 $\pm$ 0.105	18.413 $\pm$ 0.105	-0.104 $\pm$ 0.032	0.56% $\pm$ 0.17%

^aNumber of deaths and person-years are from Reference 8.

Table 4

Hypothetical overall survival difference, cause-specific survival difference, cumulative incidence difference, and proposed estimated overall survival difference between treatment arms at five years for exponential survival with five-year cause-specific survival of 90% for the experimental treatment and 60% for the control treatment

5-year cause-specific survival for other causes of death		Difference in treatment-arm curves at 5 years			
Control	Experimental	True overall survival (OS)	Cause-specific	Cumulative incidence	Proposed OS ^(a)
50% (^b)	50% (^b)	15.0%	30.0%	22.4%	15.0%
55%	50%	12.0%	30.0%	23.6%	15.7%
60%	50%	9.0%	30.0%	24.7%	16.3%
50%	55%	19.5%	30.0%	22.1%	15.8%
50%	60%	24.0%	30.0%	21.8%	16.7%

^a Asymptotic value for proposed estimators of overall survival.

^b This scenario of equal cause-specific hazards for cause 2 corresponds to line 6 of Table 1 and is the one for which the proposed estimator was developed.