Published in final edited form as:

Clin Trials. 2016 June ; 13(3): 352–357. doi:10.1177/1740774515625990.

Assessing Survival Benefit When Treatment Delays Disease Progression

David A. Schoenfeld¹ and Dianne M. Finkelstein^{1,*}

¹Massachusetts General Hospital and Harvard University, Biostatistics Unit, 50 Staniford Street, Boston MA 02114 USA

Abstract

Background—For a potentially lethal chronic disease like cancer, it is often infeasible to compare treatments on the basis of overall survival, so a combined outcome such as progression-free survival (which is the time from randomization to progression or death) has become an acceptable primary endpoint. The rationale of using an efficacy measure that is dominated by the time to progression is that an effective treatment will delay progression and when treatment is stopped at progression, the effect of treatment after this time is small. However often trials that show a significant benefit for delaying progression but not on overall survival are not universally viewed as convincing evidence that the drug should become the standard of care.

Methods—We propose that when there is a significant treatment effect of delaying progression, a Bayesian analysis of overall survival should be undertaken. We suggest using a joint piecewise exponential model, where the treatment effect on the hazard for progression and for death after progression are captured through two distinct parameters. We develop a plot of the overall survival advantage of the new therapy versus the relative hazard for death after progression. This plot can augment the discussion about whether the new treatment is beneficial on survival.

Results—In the example of an early breast cancer trial for which a new treatment significantly delayed disease recurrence, our Bayesian analysis showed that with very reasonable assumptions on the effects of treatment after recurrence, there is a high probability that the new treatment improves overall survival.

Conclusions—For a clinical trial for which treatment delays progression, the proposed method can improve the interpretability of the survival comparison using data from the study.

Keywords

Progression-free survival (PFS); Disease-free survival (DFS); Piecewise exponential; Survival; Interval-censored data; Bayesian; Cancer; Clinical Trial

*Corresponding Author: dfinkelstein@mgh.harvard.edu.

Trial Registration: clinicaltrials.gov NCT00003140

Introduction

For a potentially lethal chronic disease like cancer, the clearest sign of benefit of a new therapy is a clinically meaningful and statistically significant delay in mortality. However, in many studies of early stage cancer, the low disease-specific mortality rate makes it infeasible to perform a trial that will show such an advantage. In light of this, many clinical trials today are designed to compare therapies on the basis of the progression-free survival time, which is the time from randomization until the first of confirmed disease progression or death. In the settings where ablative surgery is possible, disease-free survival (DFS, which is time to recurrence or death) or event-free survival (EFS) are used as the primary outcome. For this paper we will primarily use the term PFS, but DFS could be used interchangeably.

For cancer trials, patients are monitored at regular visits and evaluated for evidence of progression, which is an increase in the size of a previously treated lesion or evidence of a new site of disease. Since progression usually precedes cancer death (although it is not always identified before the death), the endpoint of progression-free survival (PFS) can have more power to detect the advantage of a new therapy than overall survival would have. In addition, because patients who progress often change therapy at progression, the endpoint of PFS is usually less contaminated by the effects of such salvage therapies. Finally, disease progression is usually associated with significant morbidity, and thus evidence of treatment benefit on PFS is viewed as important to the patient even in the absence of a significant overall survival advantage. The United States Food and Drug Administration has published a guideline on when it will accept PFS as a primary endpoint in a cancer clinical trial.¹ However, using PFS as a primary outcome for a cancer trial also has challenges. Most important, a trial may show a significant treatment advantage with regard to PFS but not a significant advantage on overall survival, which can leave some regulators and/or third-party payers with an uncertainty about whether the treatment should become a standard of care.

Our goal is to develop a method to analyze survival in the setting where treatment has significantly delayed disease progression but there is not expected to be sufficient power to make a treatment comparison on the basis of overall survival. We reason that patients often stop or change treatment at progression, and that it would be useful to develop a biologically plausible model for the impact of the experimental treatment on progression as well as its impact on survival after progression. We propose a joint piecewise exponential model for time to progression and time from progression to death with single parameters for the treatment effect on each of these two functions. By using this joint model, we are able to provide a sensitivity analysis of the assumption that treatment has no impact after progression on the overall survival comparison. For this we take a Bayesian approach by putting a prior distribution on the parameter for the effect of treatment on the hazard for death after progression.

A Bayesian approach has been used in other clinical trial contexts such as medical device trials² and borrowing strength from adult trials when studying drugs used for children.³ A simpler version of this model for progression and survival was used by Berry,⁴ who reported that the power for the analysis of survival was adequate only when disease progression rapidly led to death. A Bayesian analysis of the joint distribution of progression and death

was proposed in recent paper,⁵ for the goal of improving efficiency for testing survival by including auxiliary information on progression.

We will first describe the model and show how it can be used to make inferences on the treatment effect on the hazard of progression (equivalent to PFS) and how it can be used to calculate an interval estimate for the treatment effect on the probability of survival past a specified time. We present an example of the use of this model in a cancer clinical trial that has previously been reported.

Methods

Notation and Model

Suppose that a patient is last known to be free of progression at time t_l (left endpoint of censoring interval), first known to have progressed at time t_r (right endpoint of censoring interval), and is known to have died (or ended follow-up) at time t_d . Thus we have interval censored progression time data. Often the progression time is treated as exact (at t_l) or right censored. However, we assume progression occurred between t_l and t_r . If a patient has no progression time recorded, but dies at t_d , we assume they progressed at or before t_d . In order to have a likelihood which is in closed form we assume that hazard functions for time to progression and time from progression to death are constant on a grid. Thus, since a constant hazard in each time interval implies exponentially distributed survival times with different means in each interval, we are applying piecewise exponential models on time to progression and on survival after progression. If we knew both the time of progression, t_p and the time of death t_d exactly, the calculation of the likelihood would be straight-forward. The complication is that we do not know t_p exactly, and thus the likelihood in these two hazards (for progression and for death given progression) get entwined.

To simplify the notation we describe the model and the likelihood for a single patient without baseline covariates or treatment. Then we show how these are readily incorporated in the likelihood. Denote the log hazard for progression in the j^{th} interval (at a time after time $j-1$ and before or at time j) as $\lambda_p(j)$, and for a patient who progressed at the j^{th} time, the log hazard for death in the k^{th} interval is $\lambda_d(k|j)$.

Time is divided into m equal intervals, and thus the underlying log hazard for progression is described by m parameters $\lambda_p(j)$, for $j = 1, \dots, m$. To simplify the discussion assume that time has been rescaled so that the maximum recorded time is m units so that the intervals are $(0, 1)$, $(1, 2)$, $\dots$, $(m-1, m)$. Since the time range where patients are under observation is divided into m intervals, there are $m(m+1)/2$ parameters, say $\lambda_d(k|j)$, $j = 1, \dots, m$, $k = j, \dots, m$ that define the underlying hazard of death at k after progression at j . In Figure 1, we depict these on a grid, where the pair of observations on progression (shown as an interval with endpoints t_l and t_r on the x axis) and death (shown as time u on the y axis). Similarly define $\Lambda_p(t)$ as the cumulative hazard of progression at time t and $\Lambda_d(u|t)$ as the cumulative hazard of death at time u for a patient who progressed at time t .

To develop the likelihood, consider the two cases for data observed on patients. Some patients will be lost to follow up before progression has occurred. In that case, the last time they were observed (free of disease) is t_l and their contribution to the likelihood will be

$$\exp\{-\Lambda_p(t_l)\}.$$

Otherwise, let $[x]$, denote the greatest integer $\leq x$. If the time of progression is known exactly at t and the indicator for death is δ , and u is either the death time or the last time observed, then their likelihood contribution is

$$\exp\{\delta\lambda_d([u]+1|[t]+1)+\lambda_p([t]+1)-\Lambda_d(u|t)-\Lambda_p(t)\}. \quad (1)$$

Note that we require indices of the form $[t] + 1$ because $[0, 1)$ is interval 1, $[1, 2)$ is interval 2, etc.

We assume the data on progression is *interval censored*, i.e. all that is known is that the progression time occurred between the previous visit at which the patient was disease free and the next visit when new disease was detected. Thus the data on progression time consist of overlapping intervals of failure as well as right censored cases (for patients who at the end of the trial were still alive and disease free), and standard methods such as the proportional hazards model cannot be directly applied. A piecewise exponential model easily incorporate interval censoring because the likelihood can be computed analytically. In fact, when the progression time is censored into the interval between t_l and t_r the likelihood is the integral of (1) over the interval from t_l to t_r . In Appendix Section A we calculate the closed form for this integral to derive the maximum likelihood estimates for our model.

To include covariates in the model we note that if we let x_p be a vector of covariates that affect the time to progression and x_d be covariates that affect the time from progression to death and both sets would include a treatment indicator, then we assume that the hazard for progression for an individual is multiplied by $\exp(\theta_p x_p)$ and the hazard for death is multiplied by $\exp(\theta_d x_d)$. This is also described in the Appendix section A. The procedures described in this paper are available in an R package `pssm` available from the Comprehensive R Archive.

Assessing the Treatment Effect

There are two parameters in the model for assessing the treatment effect: the treatment effect on the time to progression, which is a component of θ_p and the treatment effect on the time from progression to death which is a component of θ_d . Both of these are log-hazard ratios and can be thought of as approximately the % change in the hazard of progression and the % change in the hazard of death after progression. For the moment, assume that there are no covariates other than treatment. The primary analysis of a cancer study should test the null hypothesis that $\theta_p = 0$. This tests whether the new treatment delays progression. We assume that if death occurs before progression is noted, then progression has occurred between the

last disease free visit and the time of death. Death without progression is not explicitly incorporated in this model. If a death were clearly unrelated to the cancer, then one could simply censor the patient at this point in time. On the other hand, if one needed to include such deaths as part of the effect of treatment, one would need to model this as a competing risk process.

The effect of a treatment on overall survival is not captured in the single parameter θ_d but is effected by θ_p and other model parameters (see Appendix B). In fact if θ_d were known to be zero then overall survival would improve whenever time to progression improved, because given no treatment effect on subsequent survival, patients on the active treatment arm who progressed later would then survive longer than patients (on control arm) who progressed earlier.

If we formulate the measure of overall survival to be the difference in the survival curves at a specified time (determined by the designed follow-up time of the trial), we can calculate the upper $\alpha\%$ probability boundary on the difference between survival on the placebo and survival on the active therapy. If the upper bound were zero, it would mean that the survival of treated patients was at least as good or better than survival of the control group patients. The problem is that this bound may be too high to be useful. This is especially true in clinical trials that were stopped early due to superiority based on progression-free survival. In these trials there may be very few deaths because of inadequate follow up and overall survival differences will not be well estimated. We try to shed light on the dilemma by graphing the upper boundary of the survival difference as a function of a new parameter which measures the uncertainty about whether the treatment has any effect on the hazard of death after progression. The uncertainty is translated into a prior distribution of the log hazard ratio, a parameter that measures the effect of treatment on mortality after progression of the control group hazard divided by the active group hazard $\hat{\theta}_d$. We suppose that the mean of this distribution will be zero and that uncertainty is reflected in the spread about zero, or standard deviation, S_d . We would therefore like to posit a value of S_d and then provide the upper bound on the survival difference. However, investigators might find the upper bound on the hazard ratio to be more understandable than the S_d and thus we instead provide a plot relating the upper bound on the log of the hazard ratio versus the upper bound on the survival difference, much as one would in a sensitivity analysis. In Appendix C we show how to calculate the probability boundary on the difference in survival curves as a function of S_d . In the example later in this paper we provide the graph relating the treatment hazard ratio on death after progression to the survival difference at a point in time.

Issues in the Interpretation of the Treatment Effect on the Time from Progression to Death

Since we are proposing that a prior be put on the time from progression to death, it is important to understand the meaning of this parameter. In the context of a cancer trial, it is reasonable to assume that treatment has little direct effect after it is halted at progression. However, there may be indirect effects of a treatment on the hazard for death after progression even when progression occurs well after the treatment has finished. For instance, the choice of second line therapies might be influenced by the earlier treatments a patient

receives. It is also possible that the initial therapy a patient receives would affect their health at the time of progression which would affect their subsequent survival.

A second factor that could impact the interpretation of θ_d is that baseline covariates that were balanced by the randomization may not be balanced at the time of progression because of differential treatment effects on subgroups of patients, and the effect of this imbalance may appear as a treatment effect on the probability of death after progression. This last problem can be partially corrected for by including baseline covariates that are predictive of time to progression in both the models for time to progression and for time from progression to death. It would seem that the use of covariates measured at the time of progression would improve the model for the time from progression to death but it would not be appropriate to use such a model to estimate the survival difference (measured from randomization) we care about because the estimates would be conditional on a future value of a covariate (at progression) which is not appropriate.⁶

When the prior distribution of θ_d is imposed, it is important to consider each of these possibilities. If there are likely to be large direct or indirect effects of initial treatment on survival after progression, then the method suggested here may not be appropriate.

Example

As an example we considered data from a large phase III placebo-controlled trial MA-17 conducted by the Canadian NCI to test Letrozole as an adjuvant therapy following surgery and tamoxifen for early stage breast cancer.⁷ The study had 5187 patients but was stopped early because of a clear superiority of Letrozole in disease-free survival. There was no significant survival difference between the arms, but there were only forty one deaths. The reduction of the hazard of recurrence was 31% with a 95% confidence interval of (13% to 46%) and p -value of 0.002. The estimate of the hazard ratio of death after recurrence was very wide from 0.65 to 2.41 but favored the placebo arm (1.25).

The overall survival curves were very similar, with 40 month (trial follow-up time) survival in both arms over 98%, where the difference in the survival curves was only 0.4% (favoring Letrozole). However, using the model without a prior distribution, the 5% upper bound on the treatment difference on survival of control minus active group patients was 0.6%, and thus one could only rule out the possibility that Letrozole was more than 0.6% worse. Figure 2 plots the upper 5% bound on the treatment survival difference (y-axis) as a function of the upper 97.5% bound on the log-hazard ratio for survival after recurrence (x-axis), based on prior for the effect of Letrozole treatment on survival after recurrence. This plot crosses zero at a log-hazard ratio of approximately .39 (hazard ratio of approximately 1.5). A bound of zero on the survival difference is equivalent to the statement that survival is significantly better on the new therapy. Thus if we were 97.5% sure that this hazard ratio for death after recurrence was no worse than 1.39, we would conclude that Letrozole improved overall survival. In this example, a fairly wide prior boundary on the effect of treatment on the hazard for death leads to the conclusion that Letrozole probably improves survival. The reason for this is that there was a large improvement in time to progression combined with a small effect of treatment following progression.

Considering the issue of the imbalance of the treatment groups at recurrence, we also did this analysis including prognostic covariates in our model. Of the patients in the Letrozole study about half had positive nodes which had a significant effect on time to recurrence ($p < .001$) and a marginally significant effect on time from recurrence to death ($p = .10$). The effect of this covariate adjustment on the treatment effect (for time from recurrence to death) was relatively minor from 0.23 (unadjusted) to 0.18 (adjusted). If we weight the strata according to their population prevalences and use the same prior distribution then the lower confidence bound becomes 0.02, which is slightly above zero (which was the unadjusted lower bound on the survival difference). In fact, to get a bound of exactly zero the 97.5% bound on the log-hazard ratio would be 0.42 (up from .39), so the inclusion of the covariates slightly improves our confidence that the survival of patients treated by Letrozole will be improved.

Discussion

These methods can be useful for interpreting the results of a trial with low mortality, but use should be limited to situations in which a trial has shown that there is a clinically meaningful and statistically significant treatment benefit on time to progression, and the issue is how to judge the value of the therapy on overall survival when the amount of data on cancer mortality is scant and is not expected to improve.

We note also that use of this approach should be limited to the situation where treatment is not expected to have negative impact on the patient after it is stopped at progression. Thus, if the treatment were known to have long-term complications (even after it is stopped), and the prior on θ_d would be wide, there would be little to gain from this method.

The analysis we propose is similar to a sensitivity analysis of the assumption that there is minimal effect of treatment after it is stopped at progression, in that we capture the uncertainty about inference on overall survival against the standard deviation of the prior distribution for θ_d . The results of the proposed methods should be viewed as potentially enhancing the understanding of the treatment impact on survival. This approach cannot replace a trial designed to allow a survival comparison. Instead, if it is not possible to complete such a trial, this method could augment the discussion of overall survival benefit.

A concern raised by reviewers of this paper was that the interpretation of the parameter θ_d is unclear because the treatment groups may not be balanced after progression. In response we suggested using baseline covariates to adjust for these differences, but it is also possible to adjust for treatment decisions at progression using methods such as dynamic treatment regimes⁸ could be applied to improve the estimate of survival after progression, allowing the prior distribution to focus on more clearly on the treatment effect, but these extensions are beyond the scope of this paper. To be most cautious, we suggest that these issues be considered in the assignment of S_d .

One of our original goals with this approach was to use progression as auxiliary information for survival, and thus by jointly modeling these outcomes, we hoped to improve the

estimation and comparison for survival. However, our simulation study on this did not find this to be a great advantage, much as the result we noted in an earlier paper on this.⁹

We chose a relatively simple model where all deaths are related to cancer. The model could be extended to a more complex model in which death without progression is modeled. The best model may depend on the treatment being studied. In some cases the treatment affect may be on intercurrent deaths and other cases this may be unlikely and requires its own prior distribution. The focus of this paper is conceptual and hopefully can guide the development of a more complex model such as one with competing risks or cure. Similarly, this paper has focused on the oncology setting, but with modification, the approach could be suitable in a wider range of chronic potentially lethal diseases.

Acknowledgments

We wish to thank Dr. Paul Goss for allowing us to use the MA17 data and Dr. Kathrin Strasser-Weippl for helpful discussions, and the editors and reviewers for thoughtful comments that revised this paper. This work was partially funded by grants from the Avon Foundation and from the NIH: CA167570, UL1RR025758, and CA160233.

Appendix

A. Derivation of the Likelihoods

To develop the likelihood, consider the cases for data observed on patients. Some patients will be lost to follow up before progression has occurred. In that case, the last time they were observed (free of disease) is t_l and their contribution to the likelihood will be

$$\exp\{-\Lambda_p(t_l)\}.$$

If the time of progression is known exactly at t and the indicator for death is δ , and u is either the death time or the last time observed, then their likelihood contribution is

$$\exp\{\delta\lambda_d([u]+1|[t]+1)+\lambda_p([t]+1)-\Lambda_d(u|t)-\Lambda_p(t)\}. \quad (1)$$

When the progression time is censored into the interval between t_l and t_r the likelihood is the product of the probability of progression at a time t and the conditional probability of dying at time u given progression at time t , and the likelihood would include the integral of this product over the interval from t_l to t_r .

In order to calculate the cumulative hazard for progression, we need to calculate the amount of time, $\tau_j(t)$, that the patient who progressed at t was in the j^{th} interval. For this, if we let $[t]$ be the greatest integer function, then $\tau_j(t) = \min(j, t) - (j - 1)$ between $j = 1$ and $j = [t] + 1$ and zero otherwise, or equivalently,

$$\tau_j(t) = 1 \text{ if } j \leq t \quad = t - (j - 1) \text{ if } j - 1 < t < j \quad = 0 \text{ otherwise}$$

Then cumulative hazard for progression at time t is given by

$$\Lambda_p(t) = \sum_{j=1}^m \exp\{\lambda_p(j)\} \cdot \tau_j(t) \quad (2)$$

Similarly, let $\phi_k(u|t)$ be the amount of time that the patient who progressed at t and died at u was in the k^{th} interval measured on the y -axis. Thus, k ranges over all possible intervals that are consistent with the time lapse from progression (which is interval censored) and death. For $k = [t] + 1$ to $k = [u] + 1$,

$$\phi_k(u|t) = (\min(k, u) - \max(k - 1, t))$$

and 0 otherwise. Then cumulative hazard for the time from progression to death is given by

$$\Lambda_d(u|t) = \sum_{k=1}^m \exp\{\lambda_d(k|[t]+1)\} \cdot \phi_k(u|t)$$

To develop the likelihood, consider the cases for data observed on patients. Some patients will be lost to follow up before progression has occurred. In that case, the last time they were observed (free of disease) is t_l and their contribution to the likelihood will be

$$\exp\{-\Lambda_p(t_l)\}.$$

If the time of progression is known exactly at t and the indicator for death is δ , and u is either the death time or the last time observed, then their likelihood contribution is

$$\exp\{\delta\lambda_d([u]+1|[t]+1) + \lambda_p([t]+1) - \Lambda_d(u|t) - \Lambda_p(t)\}. \quad (3)$$

If we observe the death time t_d and the interval into which t is censored, $(t_l, t_r]$, then the likelihood in (3) will include an integral between t_l and t_r . We derive the closed form for this integral as follows by first noting that the likelihood is given by:

$$\sum_{i=[t_l]+1}^{[t_r]+1} \int_{\max(i-1, t_l)}^{\min(i, t_r)} \exp\left\{\delta\lambda_d([t_d]+1|i) + \lambda_p(i) - \sum_{j=i}^{[t_d]+1} \exp(\lambda_d(j|i))\phi_j(t_d|t) - \sum_{j=1}^i \exp(\lambda_p(j))\tau_j(t)\right\} dt. \quad (4)$$

For observations for whom the progression time is known exactly, the likelihood contribution would have the same expression as above with the outside summation and integral removed.

We note that the interval-censored nature of the data on progression time result in a likelihood which contains an integral of the probability of progression over the interval where it could have occurred. However, a piecewise exponential model more easily incorporates interval censoring because this integral can be computed analytically. In general, the integral in (4) can be removed by noting that the first term of the expression does not involve t , and the second two terms can be expanded so terms in t are isolated. In fact these two terms are

$$- \sum_{j=i+1}^{[t_d]+1} \exp \lambda_d(j|i) [\min(j, t_d) - (j-1)] - \exp \lambda_d(i|i) [\min(i, t_d) - t] \\ - \sum_{j=1}^{i-1} \exp(\lambda_p(j)) - \exp(\lambda_p(i)) [t - (i-1)]$$

Thus the coefficient of t in this equation is:

$$h_i = \exp\{\lambda_d(i|i)\} - \exp\{\lambda_p(i)\}$$

We note that this portion of the likelihood comes from the squares of the grid in Figure 1 that occur along the diagonal $k=j$, which are the triangles of the shaded region which change as t moves between t_l and t_r . Thus, the cumulative hazard is constant in the rectangles, and the effect of t moving between t_l and t_r is only on the amount of the shaded region in the triangle along $k=j$.

Then the likelihood can be written

$$= \sum_{i=[t_l]+1}^{[t_r]+1} \exp [\delta \lambda_d([t_d] \\ + 1|i) \\ + \lambda_p(i) \\ - \sum_{j=i+1}^{[t_d]+1} \exp(\lambda_d(j|i)) (\min(j, t_d) \\ - (j-1)) \\ - \sum_{j=1}^{i-1} \exp(\lambda_p(j)) \\ + \exp(\lambda_p(i)) (i \\ - 1) - \exp(\lambda_d(i|i)) \min(i, t_d)] \quad (5)$$

$$\times \frac{\exp(h_i \min(i, t_r)) - \exp(h_i \max(i - 1, t_l))}{h_i} \quad (6)$$

Covariates

To include covariates in the model we note that if we let x_p be a vector of covariates that affect the time to progression and x_d be covariates that affect the time from progression to death; both sets would include a treatment indicator. Bringing these covariates into the model, we assume that the hazard for progression for an individual is multiplied by $\exp(\theta_p x_p)$ and the hazard for death is multiplied by $\exp(\theta_d x_d)$. The derivation of the likelihood, and consequently the computer algorithm, can be simplified by adding $\theta_p x_p$ to $\lambda_p(j)$ and adding $\theta_d x_d$ to $\lambda_d(k|j)$ before deriving the likelihood. Then terms in x_p and x_d don't appear in the likelihood. In what follows, to simplify the exposition, we assume that x_d and x_p are both simply the treatment indicator, but more complicated models can be accommodated easily. The procedures described in this paper are available in an R package `pssm` available from the Comprehensive R Archive.

Computational Issues

The maximum likelihood estimates for θ_p , θ_d and all of the $m + \frac{1}{2}m(m+1)$ hazards associated with progression $\lambda_p(j)$ and death following progression $\lambda_d(k|j)$ respectively, for $j = 1 \dots m$ and $k = 1 \dots m$ are estimated using an optimization routine. Accomplishing this requires overcoming several issues related to the complexity and number of parameters involved. First, the expression in (6) is the difference of two large quantities, resulting in instability due to roundoff noise in the difference. To prevent numerical underflow problems, we suggest the following: let the terms in the sum in (5) to be denoted as I_i . By adding the exponent in I_i to each of the two terms in h_i in (6) before exponentiating and subtracting, the underflow is avoided. The second issue is that if $h_i = 0$, the expression in (6) blows up. Thus, it is useful to test whether h_i is zero and in that case replace (6) by the limit which is $\min(i, t_r) - \max(i - 1, t_l)$. In addition if $t_r = t_l$ replace (6) by $\exp(h_i t)$. Third, the expression for the likelihood is very complex and high dimensional. However, it can be readily converted into an expression involving products of $n \times m$ arrays, where n is the number of patients, and this speeds up the computation in an interpreted language such as R. Fourth, to obtain the maximum the R language offers the option of whether or not to calculate derivatives. However, calculating derivatives takes longer to run than optimizations that use numerical derivatives. The reason for this is that the likelihood calculation reduces to a sequence of a few matrix multiplications that are performed very quickly in R while the calculation of the derivatives requires the use of the *apply* function across another dimension.

B. Estimating the Time to Progression and Death

Once the MLE are found, the estimates for the progression and survival curves are straightforward. We note that the probability of progressing after time t is the same as the

likelihood function for a patient who hasn't progressed the last time they were observed at time t , $\mathcal{L}(\delta = 0, t_j = 0, t_r > t)$ from the formula 2 and 3 above

$$\exp[-\hat{\Lambda}_p(t)] = \exp\left[-\sum_{j=1}^l t_j \exp(\hat{\lambda}_p(j)) \tau_j(t)\right]$$

Similarly, the probability of surviving greater than time t is the sum of two probabilities. The first would be the likelihood for a patient who hasn't progressed the last time they were observed at time t (as above) and the second would be the same as the likelihood of a patient who's last time observed before progression was $t_j = 0$ and who was observed to have progressed before t but was not observed to have died. This can be found using the formula for the likelihood, $\mathcal{L}(\delta = 0, t_j = 0, t_r = t, t_d = t)$.

The variance covariance-matrix of the progression and survival curves is readily calculated by calculating the gradient and Hessian of the relevant function above, say $\mathcal{L}$ and V numerically and then calculating the variance as $\mathcal{L}' V \mathcal{L}$.

C. Analyzing the Treatment Effect on Progression and Survival

Let θ_d be the parameter in the model which is the log-hazard ratio of the effect of treatment on the time from progression to death. Our basic idea is to assume an informative normal prior on θ_d and a non-informative prior on all the other parameters in the model. In most cases where treatment stops at progression this prior will be centered at zero and will have finite precision, p , which must be elicited. We assume a normal prior with zero precision on all the other parameters. To obtain the mean and variance of the estimate, assume that the θ_d is the last parameter, say n^{th} , in the vector of parameters and D is a matrix which is zero everywhere except for $D_{m,m}$ where it has a value p which is the prior precision of θ_d . Let V be the variance covariance matrix of the parameters in the model. Then if the prior mean of θ_d is zero, the posterior distribution of the parameter vector has a mean of $(D + V^{-1})^{-1} V^{-1} \hat{\theta}$ and variance $(D + V^{-1})^{-1}$. Note that $(D + V^{-1})^{-1} = V - V D V / (1 + p \sigma_{m,m}^2)$, so this calculation can be accomplished without inverting V . These estimates of the model, which have been updated using the prior belief about θ_d can be used to determine the distribution of the difference in survival between the two treatment groups at a given point in time. This allows a statement about the upper bound on this difference, which can be used in an assessment of overall survival from the trial.

Thus the procedure is to choose p , estimate the parameter vector $\hat{\theta} = (\hat{\theta}_p, \hat{\theta}_d)$, transform it as above and calculate the difference in the survival curves at time t as above for control minus treatment then divide by the standard error as calculated using the variance covariance matrix described above. This is used to get the confidence bounds on survival benefit. A computer program for doing this estimation `pssm` is available in the R-archives. The program estimates the parameters using the R function `optim` to maximize the likelihood. If convergence does not occur it reduces the number of intervals by one and repeats the

optimization. It would probably be possible to use a MCMC algorithm to maximize determine the posterior distribution of the parameters.

References

1. Guidance for Industry Clinical Trail Endpoints the Approval of Cancer Drugs and Biologics. 2007. <http://www.fda.gov/downloads/Drugs/.../Guidances/ucm071590.pdf>
2. Berry, D. [accessed 1 December 2015] Using a Bayesian Approach in Medical Device Development. <ftp://ftp.isds.duke.edu/pub/WorkingPapers/97-21.ps>
3. Schoenfeld DA, Zheng H, Finkelstein DM. Bayesian design using adult data to augment pediatric trials. *Clin Trials*. 2009; 6(4):297–304. [PubMed: 19667026]
4. Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst*. 2009; 101:1642–49. [PubMed: 19903805]
5. Conlon ASC, Taylor JMG, Sargent DJ. Improving efficiency in clinical trials using auxiliary information: Application of a multi-state cure model. *Biometrics*. 2015; 71:460–468. [PubMed: 25585942]
6. Kalbfleisch, JD.; Prentice, RL. *The Statistical Analysis of Failure Time Data*. Second. West Sussex, England: Wiley; 2002.
7. Goss PE, Ingle JN, Martino DO, et al. A Randomized Trial of Letrozole in Post-menopausal Women after Five Years of Tamoxifen Therapy for Early-Stage Breast Cancer. *N Engl J Med*. 2003; 349:1793–1802. [PubMed: 14551341]
8. Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of Dynamic Treatment Regimes via Inverse Probability Weighting. *Basic Clin Pharmacol Toxicol*. 2006; 98(3):237242.
9. Finkelstein DM, Schoenfeld DA. Analyzing survival in the presence of an auxiliary variable. *Stat Med*. 1994; 13:1747–54. [PubMed: 7997708]

Figure 1. Data Grid For Time to Progression and Time from Progression to Death

Figure 2.

Plot of Upper 5% Bound of Survival Difference (Placebo-Drug) versus upper 2.5% Bound on Log-hazard Ratio for Recurrence to Death (RD) of Placebo/Drug. The box represents the region where the drug is better than the placebo on survival.