Journal of Applied Statistics
Vol. 00, No. 00, Month 200x, 1–12

RESEARCH ARTICLE

Detecting Multiple Change Points in Piecewise Constant Hazard Functions

Melody S. Goodman [a*], Yi Li[b] and Ram C. Tiwari [c]

[a]*Graduate Program in Public Health, Department of Preventive Medicine, Stony Brook University School of Medicine, Stony Brook, NY, U.S.A.* [b]*Department of Biostatistics, Harvard University and Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute Boston, MA, U.S.A.* [c]*Office of Biostatistics, Center for Drug Evaluation & Research, Food and Drug Administration Silver Spring, MD, U.S.A.*

(*Received 00 Month 200x; in final form 00 Month 200x*)

The National Cancer Institute (NCI) suggests a sudden reduction in prostate cancer mortality rates, likely due to highly successful treatments and screening methods for early diagnosis. We are interested in understanding the impact of medical breakthroughs, treatments, or interventions, on the survival experience for a population. For this purpose, estimating the underlying hazard function, with possible time change points, would be of substantial interest, as it will provide a general picture of the survival trend and when this trend is disrupted. Increasing attention has been given to testing the assumption of a constant failure rate against a failure rate that changes at a single point in time. We expand the set of alternatives to allow for the consideration of multiple change-points, and propose a model selection algorithm using sequential testing for the piecewise constant hazard model. These methods are data driven and allow us to estimate not only the number of change points in the hazard function but where those changes occur. Such an analysis allows for better understanding of how changing medical practice affects the survival experience for a patient population. We test for change points in prostate cancer mortality rates using the NCI Surveillance, Epidemiology,and End Results dataset.

Keywords: survival analysis; change points; piecewise constant; hazard function; cancer

1. Introduction

According to the National Cancer Institute there has been a sudden reduction in mortality rates for prostate cancer, likely due to highly successful treatments, screening methods for early diagnosis, and public health programs [1]. This reduction in mortality rates has sparked interest towards a better understanding of the impact of medical breakthroughs, treatments, and interventions, on the survival experience for a population. We are interested in estimating the underlying hazard function, with possible time change points, as it will provide a general picture of the survival trend and when this trend is disrupted.

An estimate of the overall survival trend for an entire patient population can provide a better understanding of how changing medical practice, health policy, the role of the health care system, and public health services, affect the population survival experience. Furthermore, from an individual perspective, understanding the hazard function including how and when changes in the risk for mortality occur, allows for a life course view.

Several authors (see for example, [2, 10, 11]) have proposed methods for detecting a single change point in a piecewise constant hazard function. However, with advances in

*Corresponding author. Email: melody.goodman@gmail.com

ISSN: 0266-4763 print/ISSN 1360-0532 online

DOI: 10.1080/0266476YYxxxxxxxx
http://www.informaworld.com

biomedical research and increasing survival rates of patients we believe there may be some cases where a model with two or more change points is more appropriate [8]. We consider the situation where there are multiple changes in the piecewise constant hazard function at a few unknown time points, often referred to as the change points. We propose a data-driven approach for detecting the number of change points in the hazard function, and estimating all of the unknown parameters in the model, including the change points. An alpha-spending scheme [9] is developed such that the maximum number of change points, K, does not need to be prespecified. We propose a Wald type test statistic based on the theoretical results of Yao [17], as the likelihood ratio test is controversial in a change-point setting due to the irregularity of models and issues of singularity [4, 11, 12, 15]; simulation studies confirm the validity of our proposed test.

The rest of this article is structured as follows. We define the piecewise constant model with multiple change points and develop a Wald type test statistic in Section 2. We discuss the issue of sequential testing and model selection using an alpha-spending type function in Section 3. For methodology assessment, simulation studies to investigate the overall Type I error, power, and estimation of parameters are presented in Section 4. We apply our proposed methods to test for change points in the mortality rates of prostate cancer patients in the NCI Surveillance, Epidemiology, and End Results (SEER) Program in Section 5. We conclude with a general discussion in Section 6.

2. Piecewise Constant Multiple Change Point Model

Let $X_1,...,X_n$ denote independent identically distributed survival times, and $C_1,...,C_n$ be the censoring times which are assumed to be independent of X. We only observe the pairs $(T_i,\delta_i), i = 1,2,...,n$, where $T_i = \min(X_i,C_i)$ and $\delta_i = 1$ if $X_i \leq C_i$ and zero otherwise. Consider the following change point model:

$$\lambda(t) = \begin{cases} \alpha_1 & 0 \leq t \leq \tau_1 \\ \alpha_2 & \tau_1 < t \leq \tau_2 \\ \vdots & \\ \alpha_{k+1} & t > \tau_k, \end{cases} \tag{1}$$

where $0 = \tau_0 < \tau_1 < \ldots < \tau_k = \infty$ are the change points, k is the number of change points in the model, and α_j is the value of the hazard function between the time points τ_{j-1} and τ_j.

We propose maximum (profile) likelihood estimation to estimate the unknown parameters of α and τ. Based on equation 1, the log-likelihood function is

$$logL(\alpha_1,...,\alpha_{k+1},\tau_1,...,\tau_k) = \sum_{j=1}^{k+1}[X(\tau_j) - X(\tau_{j-1})]\log\alpha_j - \sum_{i=1}^{n}\sum_{j=1}^{k+1}\alpha_j(T_i \wedge \tau_j - T_i \wedge \tau_{j-1})$$

where $X(t) = \sum_{i=1}^{n} I(T_i \leq t, \delta_i = 1)$ is the number of deaths observed up to time t. With $\tau_j, j = 1,\ldots,k$, fixed, some algebra yields that the maximizers of $\alpha_j, j = 1,\ldots,k+1$ are given by $\hat{\alpha}_j = \frac{X(\tau_j)-X(\tau_{j-1})}{\sum_{i=1}^{n}(T_i \wedge \tau_j - \tau_{j-1})I(T_i > \tau_{j-1})}$. Substituting these values into $logL$ gives the profile likelihood for τ_j's. Which can be expressed as,

$$\ell(\tau_1,...,\tau_k) = \sum_{j=1}^{k+1}\{X(\tau_j) - X(\tau_{j-1})\}\log\left\{\frac{X(\tau_j) - X(\tau_{j-1})}{\sum_{i=1}^{n}(T_i \wedge \tau_j - \tau_{j-1})I(T_i > \tau_{j-1})}\right\}.$$

We then maximize $\ell(\tau_1, ..., \tau_k)$ with respect to $\tau_j, j = 1, \ldots, k$ and insert the obtained values back to $\hat{\alpha}_j, j = 1, \ldots, k+1$ for the MLEs of α_j.

Our primary interest is to test for the existence of change points in the hazard function; or equivalently, to test whether the hazard before and after such points are different. To test the null hypothesis of no change points against the alternative of one change point, we develop a hypothesis test based on the hazard parameters ($H_0 : \alpha_1 - \alpha_2 = 0$). Yao [17] showed that the estimates of the α_j and τ_j are independent. Based on this, we formulate the Wald type test statistic solely on the hazard rate (α_j) parameters while treating the change point parameters as fixed.

Hence, to test $H_0 : \alpha_{k-1} - \alpha_k = 0$ versus $H_1 : \alpha_{k-1} - \alpha_k \neq 0$, we use a Wald type test statistic of the form,

$$X_W = \frac{(\hat{\alpha}_{k-1} - \hat{\alpha}_k)^2}{Var(\hat{\alpha}_{k-1} - \hat{\alpha}_k)}, \tag{2}$$

which asymptotically follows a Chi-squared distribution with one degree of freedom under the null hypothesis.

3. Sequential Testing and Model Selection

We assume that the maximum number of change points in the model is some finite number K. Our aim is to find the model, with the number of change points k ($k = 0, ..., K$) that best fits our data. We frame our model selection process as a sequential testing problem. We start with the model with no change points and perform a hypothesis test to compare it to the model with one change point. If we fail to reject the null hypothesis, we stop and conclude the final model has no change points. However, if we reject the null hypothesis, we will continue on to the next hypothesis test; compare the model with one change point, to the model with two change points. If the null hypothesis is rejected we test the subsequent hypothesis. That is, we will test for the existence of τ_k only after verifying the existence of τ_{k-1}. The algorithm is continued until we fail to reject a hypothesis.

We are cautious of choosing an over-fitted model with a large number of change points. The issue of selecting the appropriate number of changes points is similar in spirit to the issue of selecting bin width in a smoothing problem. To address this issue, we borrow methodology from the group sequential analysis literature. Lan and DeMets [9] proposed an alpha spending technique in which the nominal significance level needed to reject the null hypothesis at each analysis is $\leq \alpha$ and increases as the study progresses. Thus, for the test, it is more difficult to reject the null at the earliest analysis but easier later on. They proposed

$$\alpha^*(k) = \alpha s^*(k),$$

where $\alpha^*(k)$ is the significance level for the kth hypothesis and $s^*(k) = \frac{k}{K}$ is the spending function. Here, $\alpha^*(1) < \alpha^*(2) < ... < \alpha^*(K)$; an increasing alpha spending function.

In order to find a parsimonious model we want strong evidence for choosing a more complicated model, a model with more change points, over a simpler one. Therefore, we are interested in a decreasing alpha spending function, where $\alpha^*(1) > \alpha^*(2) > ... > \alpha^*(K)$; the test for each additional change point will be conducted at a more conservative α level than the one before it.

If the overall significance level is α, we propose $\alpha^*(k) = \frac{\alpha}{2^{k-1}}$, where $\alpha^*(k)$ is the significance level for the kth hypothesis test. An advantage of this alpha spending function is that it does not depend on the overall number of hypothesis tests being conducted. Therefore, one does not need to specify K, the maximum number of change-points, which

is ideal for flexibility in model selection.

In this setting, Type I error is the probability of incorrectly choosing a model that has more change points than the true model. We show that the type I error rate will not exceed α by calculating three probabilities (i) choosing a model with one or more change points given the true model has no change points, (ii) choosing a model with k change points given the true model has no change points, and (iii) choosing a model with more than k change points given the true model has k change points, are all at most α.

Let M_i be the event that the model has i change points. Then

$$P(M_{i\geq 1}|M_0) = 1 - P(M_0|M_0) = 1 - (1-\alpha) = \alpha, \tag{3}$$

$$P(M_k|M_0) = \left(1 - \frac{\alpha}{2^k}\right)\prod_{i=1}^{k}\frac{\alpha}{2^{i-1}} < \frac{\alpha^k}{\prod_{i=1}^{k} 2^{i-1}} < \alpha, \tag{4}$$

and for $k \geq 1$

$$P(M_{i>k}|M_k) = \sum_{j=1}^{\infty}\left[\left(1 - \frac{\alpha}{2^{k+j}}\right)\prod_{i=k+1}^{k+j}\frac{\alpha}{2^{i-1}}\right] < \alpha. \tag{5}$$

The derivations of (4) and (5) can be found in Appendix A.

4. Simulation

In order to test our proposed methodology we conducted simulation studies on models with two, three, and four change points. In this paper we discuss and present results for the model with two change points as this is the simplest multiple change point case and simulation studies for models with a greater number of change points are easily extrapolated from those of two change point model. We conducted simulation studies to investigate the estimation of model parameters, the power of our proposed test, and the overall Type I error rate.

Survival times for a piecewise constant hazard model with two change points were generated by inverting the cumulative distribution function and using the probability integral transformation. In this case three different distributions have to be generated; (1) survival times before the first change points, (2) survival times between the first change point and the second change point and (3) survival times after the second change point. Censoring times were generated from three corresponding uniform distributions. Five parameters need to be specified a priori: the value of the hazard function for the three distributions and the two change points. We developed simulation scenarios that closely approximated those of the SEER data from the application in Section 5 and simulated models under the following conditions $0 < \alpha_k < 1$, $0 < \tau_1 < 5$, and $\tau_1 < \tau_2 < 10$. Censoring percentages were calculated for each simulated data set. We manipulated the amount of censoring by changing the ranges on the censoring distributions. We tested a variety of model scenarios and compared the average of each simulated data set to the theoretical average of the specified model.

5,000 independent data sets were created, for each data set we use the Wald type test statistic (2) to compare the null model with no change points to an alternative model with one change point. If we reject the null we continue to the next hypothesis and test the null model of one change point to the alternative model of two change points. We continue this algorithm until we are unable to reject a hypothesis. The α level for each test was determined using the decreasing alpha spending function proposed in Section 3. For

each simulated dataset a final model was chosen and the parameters were estimated by maximizing the likelihood function for the final model using the Nelder-Mead Simplex algorithm, a reliable procedure to carry out such optimization (available in common statistical software). Table 1 displays the results of our simulations for $n = 500$, under 7 different model scenarios:

(1) $\alpha_1 = 0.95$, $\alpha_2 = 0.55$, $\alpha_3 = 0.15$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 0% censoring;
(2) $\alpha_1 = 0.15$, $\alpha_2 = 0.55$, $\alpha_3 = 0.35$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 5% censoring;
(3) $\alpha_1 = 0.15$, $\alpha_2 = 0.35$, $\alpha_3 = 0.55$, $\tau_1 = 1.0$, $\tau_2 = 3.5$, and 20% censoring;
(4) $\alpha_1 = 0.65$, $\alpha_2 = 0.35$, $\alpha_3 = 0.15$, $\tau_1 = 1.5$, $\tau_2 = 3.5$, and 25% censoring;
(5) $\alpha_1 = 0.15$, $\alpha_2 = 0.35$, $\alpha_3 = 0.55$, $\tau_1 = 1.0$, $\tau_2 = 2.5$, and 30% censoring;
(6) $\alpha_1 = 0.15$, $\alpha_2 = 0.55$, $\alpha_3 = 0.95$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 35% censoring;
(7) $\alpha_1 = 0.15$, $\alpha_2 = 0.95$, $\alpha_3 = 0.45$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 50% censoring.

The mean estimated value is the average estimated parameter value from all 5,000 simulation runs, and the standard error is the standard deviation of these estimates. Based on simulation studies our method estimates the change points and the value of the hazard quite well even with a moderate amount of censoring. In most cases the coverage probability is close to the nominal level of 95%. However, as the percentage of censoring increases above 40% the coverage probability decreases significantly (see Table 1).

We examined the power of the test under the two change points alternative. The power is essentially a measure of the accuracy of our algorithm in choosing the final model. Table 2 displays the results of the power analysis for samples of size 500 under 10 different model scenarios:

(1) $\alpha_1 = 0.15$, $\alpha_2 = 0.55$, $\alpha_3 = 0.95$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 1% censoring;
(2) $\alpha_1 = 0.15$, $\alpha_2 = 0.55$, $\alpha_3 = 0.35$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 3% censoring;
(3) $\alpha_1 = 0.95$, $\alpha_2 = 0.55$, $\alpha_3 = 0.15$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 16% censoring;
(4) $\alpha_1 = 0.15$, $\alpha_2 = 0.35$, $\alpha_3 = 0.55$, $\tau_1 = 1.0$, $\tau_2 = 3.5$, and 23% censoring;
(5) $\alpha_1 = 0.95$, $\alpha_2 = 0.35$, $\alpha_3 = 0.75$, $\tau_1 = 1.5$, $\tau_2 = 3.5$, and 26% censoring;
(6) $\alpha_1 = 0.95$, $\alpha_2 = 0.65$, $\alpha_3 = 0.45$, $\tau_1 = 1.0$, $\tau_2 = 3.5$, and 27% censoring;
(7) $\alpha_1 = 0.15$, $\alpha_2 = 0.35$, $\alpha_3 = 0.55$, $\tau_1 = 1.0$, $\tau_2 = 3.5$, and 27% censoring;
(8) $\alpha_1 = 0.15$, $\alpha_2 = 0.55$, $\alpha_3 = 0.95$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 36% censoring;
(9) $\alpha_1 = 0.15$, $\alpha_2 = 0.95$, $\alpha_3 = 0.45$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 49% censoring;
(10) $\alpha_1 = 0.15$, $\alpha_2 = 0.95$, $\alpha_3 = 0.45$, $\tau_1 = 2.0$, $\tau_2 = 4.0$, and 57% censoring.

Power was most affected by sample size, the difference between α_{k-1} and α_k, and the difference between τ_{k-1} and τ_k. With samples of size 500, $|\alpha_{k-1} - \alpha_k| > 0.2$, $|\tau_{k-1} - \tau_k| > 1.5$, and $< 30\%$ censoring, we observed power of at least 90%, however as the censoring percentage increase above 30% the power decreases (see Table 2).

To test the overall Type I error rate, we simulated data with two change points ($\alpha_1 = 0.95$, $\alpha_2 = 0.55$, $\alpha_3 = 0.25$, $\tau_1 = 2$, and $\tau_2 = 4$) and implemented our stepwise model selection algorithm. We developed a censoring distribution using three uniform distributions, zero to the first change point, from the first change point to the second change point, and from the second change point to the largest survival time. On average this distribution produces a 35% censoring percentage; we decrease the censoring percentage by increasing the upper limits on the censoring distributions and increase the censoring percentage by decreasing the upper limits on the censoring distributions. Table 3 displays a sample of our results for samples of size 500. The Type I error is the number of times our algorithm chooses a final model with three or more change points over the null model of two change points. The observed overall Type I error rate is similar to the nominal significance level of the test (see Equation 5).

5. Prostate Cancer Mortality

To examine prostate cancer mortality, we analyze the Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) Public-Use Data (1973-2007), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch. This data set contains cancer incidence and survival for cases diagnosed from 1973 to 2007, follow-up continued until December 31, 2007. We are interested in finding: (i) if there are change points in the hazard function, (ii) the number of change points and (iii) estimating the location of the change points and all other model parameters.

We define an event as death from prostate cancer. If a subject dies from another cause they are censored at the time of their death. We restrict our analysis to men who were diagnosed with prostate cancer between 1973 and 2002 to allow for at least five years of follow-up. For the purpose of analysis we excluded subjects with unknown follow-up time. To avoid singularity, we restrict the change points to be larger than the first survival time and smaller than the second to last survival time, assuming these are non-censored time points, $T_{(1)} < \tau_1 < \ldots < \tau_k < T_{(n-1)}$ ([17]; [13]).

There were 378,095 men who fit our restriction criteria in the SEER data set with 75,343 events and 80.1% of the observations being censored. The average age of men in the sample is 84 with an average age at diagnosis of 71. In our sample population 84% of the men are non-Hispanic white, 11% are non-Hispanic Black and approximately 5% are another race/ethnicity (i.e., Hispanic, American Indian/Alaska Native or Asian). The estimated hazard function has two change points (namely, at 3.0 and 5.4) and is defined by Equation 6.

$$\lambda(t) = \begin{cases} 0.0334 & 0 \leq t \leq 3.0 \\ 0.0249 & 3.0 < t \leq 5.4 \\ 0.0216 & t > 5.4. \end{cases} \tag{6}$$

Before the first change point there are 32,247 events and 63% censoring, between the first change point and the second change point there are 16,140 events and 71% censoring and after the second change point there are 26,956 events and 89% censoring. Figure 1 displays the estimated hazard function. The hazard function for men diagnosed with prostate cancer between 1973 and 2002 starts at 33.4 per 1000 population until 3 years after diagnosis, it decreases to 24.9 per 1000 population until 5.4 years after diagnosis and then decreases to 21.6 per 1000 population until the end of the follow-up period. When a one change point model is fit to this data the change point is estimated at 2.8 very close to the first estimated change point (3.0) in the two change point model.

6. Discussion

Inference about the change-point in the piecewise constant model is a nontrivial issue previously discussed in the literature (see, for example [7]; [6]; and [5]). We developed a Wald type test statistic employing the statistical independence of the estimates of the hazard rates and the change point parameters ([17]). The Wald test has a known asymptotic Chi-squared distribution, facilitating drawing inference. Our methodology requires use of the Nelder-Mead Simplex optimization algorithm [14], this algorithm uses only function values and is robust but can be relatively slow. It will work reasonably well for non-differentiable functions and requires the user to provide initial starting values for the parameter estimates. From our simulation studies we found these values need to be reasonable but not precise. There are other optimization methods available (e.g, quasi-Newton variable metric algorithm, conjugate gradients method, box constraints) and eas-

ily implemented in most software packages each with their own set of advantages and disadvantages [16].

An advantage of our method is its ability to appropriately handle sequential testing with careful selection of a parsimonious model over an over fitted model. Although traditional smoothing techniques (see, for example [3]) can be applied to estimate the hazard function. When using these methods the knots and the number of knots are to be pre-specified by the data analyst. Our approach uses the data to determine the number and value of the change points. In settings where there is an interest in not only knowing the trend of the hazard function but when the changes in trend occur, our proposed methodology has several advantaged when compared to existing approaches.

The methods we propose are easily implemented, using maximum likelihood estimation and Wald type test statistics, and can be applied to other important applications. The resulting estimate of the hazard function can also be used for predictive purposes and allows for non-parametric extrapolation of very long term survival by extrapolating the trend from the last change point. Change-point hazard functions may have implications in health care policy decisions. By enhancing our understanding of the changes in trends of population mortality rates; we can identify gaps, seek solutions, improve performance, and ultimately, better the public's health.

Acknowledgements

The research of Dr. Goodman was supported by National Institute of Child Health and Human Development grant 5 F31 HD043695. The research of Dr. Li was supported by National Institute of Health grant R01CA95747. Views expressed by Dr. Tiwari here are his own and do not necessarily represent those of Food and Drug Administration.

Table 1. Estimation for Piecewise Constant Model with Two Change Points Based on 5,000 simulations (n=500)

Censor	Parameters	Parameter Value	Mean Estimated Value	Std. Error Estimated Value	Model Based Std. Dev.	Coverage Probability
0%	α_1	0.95	0.953	0.053	0.046	0.949
	α_2	0.55	0.545	0.077	0.076	0.941
	α_3	0.15	0.145	0.032	0.034	0.923
	τ_1	2.0	1.992	0.118		
	τ_2	4.0	3.968	0.182		
5%	α_1	0.15	0.149	0.014	0.013	0.940
	α_2	0.55	0.558	0.039	0.035	0.947
	α_3	0.35	0.342	0.032	0.036	0.937
	τ_1	2.0	2.009	0.176		
	τ_2	4.0	4.018	0.170		
20%	α_1	0.15	0.148	0.021	0.017	0.933
	α_2	0.35	0.352	0.030	0.023	0.934
	α_3	0.55	0.579	0.059	0.060	0.951
	τ_1	1.0	1.023	0.218		
	τ_2	3.5	3.556	0.243		
25%	α_1	0.65	0.655	0.040	0.038	0.946
	α_2	0.35	0.346	0.048	0.048	0.937
	α_3	0.15	0.133	0.035	0.045	0.896
	τ_1	1.5	1.502	0.140		
	τ_2	3.5	3.486	0.208		
30%	α_1	0.15	0.148	0.019	0.019	0.935
	α_2	0.35	0.352	0.033	0.028	0.942
	α_3	0.55	0.569	0.048	0.045	0.946
	τ_1	1.0	1.012	0.092		
	τ_2	2.5	2.535	0.144		
35%	α_1	0.15	0.158	0.015	0.013	0.945
	α_2	0.55	0.550	0.051	0.046	0.943
	α_3	0.95	1.065	0.158	0.128	0.942
	τ_1	2.0	2.009	0.059		
	τ_2	4.0	4.055	0.169		
50%	α_1	0.15	0.126	0.014	0.017	0.558
	α_2	0.95	0.821	0.087	0.074	0.512
	α_3	0.45	0.387	0.071	0.084	0.874
	τ_1	2.0	2.005	0.017		
	τ_2	4.0	3.761	0.392		

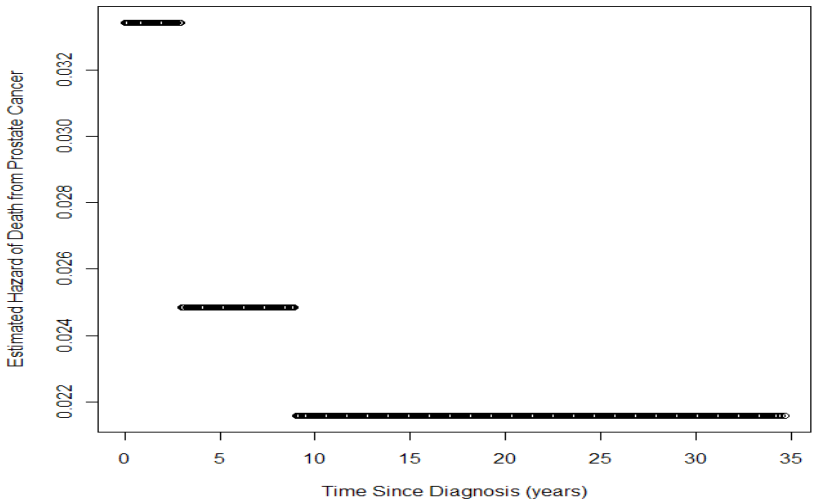
Table 2. Power Analysis for the Piecewise Constant Model with Two Change Points Based on 5,000 simulations (n=500)

$\alpha_2 - \alpha_1$	$\alpha_3 - \alpha_2$	$\tau_2 - \tau_1$	% censoring	power
0.4	0.4	2	1	0.98
0.4	-0.2	2	3	0.98
-0.4	-0.4	2	16	0.95
0.2	0.2	1.5	23	0.93
-0.6	0.4	2	26	0.92
-0.3	-0.2	1.5	27	0.93
0.2	0.2	1.5	27	0.96
0.4	0.4	2	36	0.81
0.8	-0.5	2	49	0.90
0.8	-0.5	2	57	0.69

Table 3. Type I Error Analysis for the Piecewise Constant Change-Point Model Based on 5,000 simulations (n=500)

% Censoring	Type I Error
0	0.048
5	0.048
10	0.046
15	0.044
20	0.050
25	0.052
30	0.048
40	0.049
45	0.053
50	0.052
60	0.047
80	0.053

Figure 1. **Estimated Prostate Cancer Mortality Hazard Function for Men Diagnosed 1973-2002**



References

References

[1] Edwards, B.K. and Brown, M.L. and Wingo P.A. and Howe, H.L. and Ward, E. and Ries, L.A.G. and Schrag, D. and Jamison, P.M. and Jemel, A. and Wu, X.C., and Friedman, C. and Harlan, L. and Warren, J. and Anderson, R.N. and Pickle, L.W. (2005). Annual report to the nation on the status of cancer, 1975-2002, featuring population-based trends in cancer treatment. *Journal of the National Cancer Institute* **97**, 1407–1427.

[2] Gijbels, Irène and Gürler, Ülkü (2003). Estimation of a Change Point in a Hazard Function Based on Censored Data. *Lifetime Data Analysis* **9**, 395–411.

[3] Gray, R.J. (1996). Hazard Rate Regression Using Ordinary Nonparmetric Regression Smoothers. *Journal of Computational and Graphical Statistics* **5**, 190–207.

[4] Henderson, R. (1990). A Problem with the Likelihood Ratio Test for a Change-point Hazard Rate Model. *Biometrika* **77**, 835-43.

[5] Hinkley, D.V. (1969). Inference about the intersection in two-phase regression. *Biometrika* **56**, 495-504.

[6] Hinkley, D.V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika* **57**, 1-17.

[7] Hinkley, D.V. (1971). Inference in two-phase regression. *Journal of the American Statistical Association* **66**, 736-43.

[8] Kim, Hyune-Ju, Fay, M.P., Feuer, E.J., and Midthune, D.J. (2000). Permutation Test for Joinpoint Regression with Application to Cancer Rates. *Statistics in Medicine* **19**, 335-351.

[9] Lan, K.K.G. and DeMets, D.L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* **70**,659–63.

[10] Loader, C.R. (1991). Inference for a hazard rate change point. *Biometrika* **78**, 749–57.

[11] Matthews, D. E. and Farewell, V. T. (1982). On Testing for a Constant Hazard against a Change-point Alternative. *Biometrics* **38**, 463–8.

[12] Matthews, D. E. and Farewell, V. T. (1985). On a Singularity in the Likelihood for a Change-point Hazard Rate Model. *Biometrika* **72**, 703–4.

[13] Müller, H-G. and Wang, J-L. (1994). Change-Point Models for Hazard Functions.Carlstein, E. G., Müller, H-G. and Siegmund, D. (eds) *Change-point problems*.Institute of Mathematical Statistics, 2224–40.

[14] Nelder, J.A., and Mead, R. (1965) *The Computer Journal* **7 (4)**, 308-313.

[15] Nguyen, G. S., Rogers, G.S. and Walker, E. A. (1984). Estimation in change-point hazard rate models. *Biometrika* **71**, 299–304.

[16] Nocedal,Jorge and Wright, Stephen J. *Numerical Optimization*, Springer Series in Operations Research, New York, NY, 1999.

[17] Yao, Y-C. (1986). Maximum Likelihood Estimation in Hazard Rate Models with a Change-point. *Communications in Statistics: Theory and Methods* **15**, 2455–66.

Appendix A. Proofs of Inequalities (4) and (5)

Proof of Inequality (4):

Note that

$$P(M_k|M_0) = (1-\frac{\alpha}{2^k})\prod_{i=1}^{k}\frac{\alpha}{2^{i-1}} < \frac{\alpha^k}{\prod_{i=1}^{k}2^{i-1}} < \frac{\alpha}{\prod_{i=1}^{k}2^{i-1}} < \alpha$$

where the first equality follows from the fact that $(1-\frac{\alpha}{2^k}) < 1$, the second follows from the fact that $0 < \alpha < 1$, $\alpha^k < \alpha$, and the last equality follows from the fact that $\prod_{i=1}^{k}2^{i-1} > 1$.

Proof of Inequality (5):

Note that

$$P(M_{i>k}|M_k) = \sum_{j=1}^{\infty}\left[\left(1-\frac{\alpha}{2^{k+j}}\right)\prod_{i=k+1}^{k+j}\frac{\alpha}{2^{i-1}}\right] < \sum_{j=1}^{\infty}\prod_{i=k+1}^{k+j}\frac{\alpha}{2^{i-1}} < \sum_{j=1}^{\infty}\frac{\alpha}{2^j} = \alpha$$