Journal of Modern Applied Statistical Methods

Volume 10 | Issue 2 Article 23

11-1-2011

Non-homogenous Poisson Process for Evaluating Stage I & II Ductal Breast Cancer Treatment

Chris P. Tsokos
University of South Florida, profcpt@cas.usf.edu

Yong Xu
Radford University, yxu10@radford.edu

Journal of Modern Applied Statistical Methods
November 2011, Vol. 10, No. 2, 646-655

1538 – 9472/11/$95.00

Non-homogenous Poisson Process for Evaluating Stage I & II Ductal Breast Cancer Treatment

Chris P. Tsokos
University of South Florida,
Tampa, FL

Yong Xu
Radford University,
Radford, VA

Non-Homogenous Poisson Process (NHPP), also known as the Power Law process (PLP) or the Weibull Process, is used to evaluate the effectiveness of a given treatment for Stage I & II ductal breast cancer patients. The behavior of the shape parameter of the intensity function is examined to evaluate the response of a given treatment with respect to its effectiveness for a cancer subject.

Key words: Statistical modeling, power law process, Weibull process, non-homogenous Poisson process, intensity function, cancer analysis.

Introduction

Breast cancer (malignant breast neoplasm) is cancer originating from breast tissue, most commonly from the inner lining of milk ducts or the lobules that supply the ducts with milk (Sariego, 2010). This study uses the Non-Homogenous Poisson Process (NHPP), also known as the Power Law Process (PLP) or the Weibull Process, to evaluate the effectiveness of a given treatment for Stage I & II ductal breast cancer patients. The behavior of the shape parameter of the intensity function is examined to evaluate the response of a given treatment with respect to its effectiveness for the cancer subject. Data from the Surveillance Epidemiology and End Results (SEER) Program is used to test the proposed model. This data is collected by the U.S. National Institutes of Health (NIH) (2010) and includes information on incidence, survival and prevalence from specific geographic areas representing 26% of the U.S. population; the NIH also compiles reports on several types of cancer and includes mortality rates in the SEER database.

Chris Tsokos is a Distinguished University Professor in mathematics and Statistics at the University of South Florida. His research interests are in modeling Global Warming, analysis and modeling of cancer data, parametric, Bayesian and nonparametric reliability, and stochastic systems, among others. He is a fellow of both ASA and ISI. Email him at: profcpt@cas.usf.edu. Yong Xu is an Assistant Professor in the Department of Mathematics and Statistics at Radford University. Email him at: yxu10@radford.edu.

Historical Review

Many authors have contributed to the literature on point processes. Billingsley (1961) proposed a statistical inference method for Markov processes. Duane (1964) suggested a learning curve approach to reliability monitoring. Cox and Lewis (1966) studied statistical inference problems in point processes and their applications. Cox and Isham (1980) discussed random collection of point processes, and Basawa and Parkasa Rao (1980) studied different stochastic processes with the applications. Dharmadhikari, et al. (1989) estimated the scale parameter of a power law process using power law counts. Bain and Enelhardt (1991) presented a statistical analysis of reliability and compared several life testing models. Kingman (1993) discussed methods of Poisson Sampling. Tsokos (1997) presented the parameter estimation of Power Law Process. Rigdon and Basu (2000) proposed several statistical methods for the reliability of repairable systems using a power law process.

Methodology

The schematic diagram presented in Figure 1 provides a picture of the database used in this study. A randomized data set was generated to reduce random errors by performing simple random sampling procedures. From a total 578,134 cancer patients in the SEER database, 500,000 breast cancer patients' information was randomly selected. Out of these 500,000 breast cancer patients, 496,783 are female and 3,217 are male. The female patients are categorized into three different racial groups: Caucasian, African-American and Asian (which includes others). Within these groups, there are 426,302 Caucasian, 39,681 African-American, 29,015 Asian and 1,785 unspecified patients. Within each patient group there are four types of breast cancer: ductal, medullary, lobular and other (unspecified). For each type of breast cancer, patients are further divided according to the American Joint Committee on Cancer (AJCC) Cancer Staging, such as, stage I, II, III, IV and others. Breast cancer, particularly the ductal form, is a common occurrence among Caucasian females; thus, this study focuses on ductal breast cancer among Caucasian females.

Caucasian Ductal Cancer Patients in Stage I

WD stage I stands for Caucasian ductal cancer patients in AJCC stage I. Similarly, WD stage II, III and IV stand for Caucasian ductal cancer patients in AJCC stages II, III and IV. WD patients in stage I were divided into two groups: (1) patients who are still living, and (2) patients who are deceased (see Figure 2). Deceased patients were grouped into (1) patients who are deceased due to breast cancer and, (2) patients who are deceased due to other reasons. For those patients who are deceased due to breast cancer, different treatment information is available. A NHPP was constructed with respect to WD stage I patients in order to compare the effects of the four different treatments.

Caucasian Ductal Cancer Patients in Stage II

Caucasian ductal patients in stage II were divided into two groups, patients who are still living and patients who are deceased (see Figure 3). Deceased patients were further divided into groups of patients who (1) are deceased due to breast cancer, and (2) patients who are deceased due to other reasons. For those patients who are deceased due to breast cancer, different treatment information is available. A NHPP was constructed with respect to WD stage II patients in order to compare the effects of the four different treatments.

The most common stages to classify breast cancer patients are stages I and II. Thus, these are the stages considered herein using the NHPP to determine the effectiveness of the four different treatments (see Figures 2 and 3).

Non-Homogeneous Poisson Process Analysis

According to Tsokos (1997), the non-homogeneous Poisson process (NHPP) is also known as the Power Law Process (PLP) or the Weibull process (WP), in addition, the NHPP is also considered a counting process. Let $\{N(t), t \geq 0\}$ be a counting process with the following three properties:

1. $N(t) \geq 0$.

2. $N(t)$ is an integer.

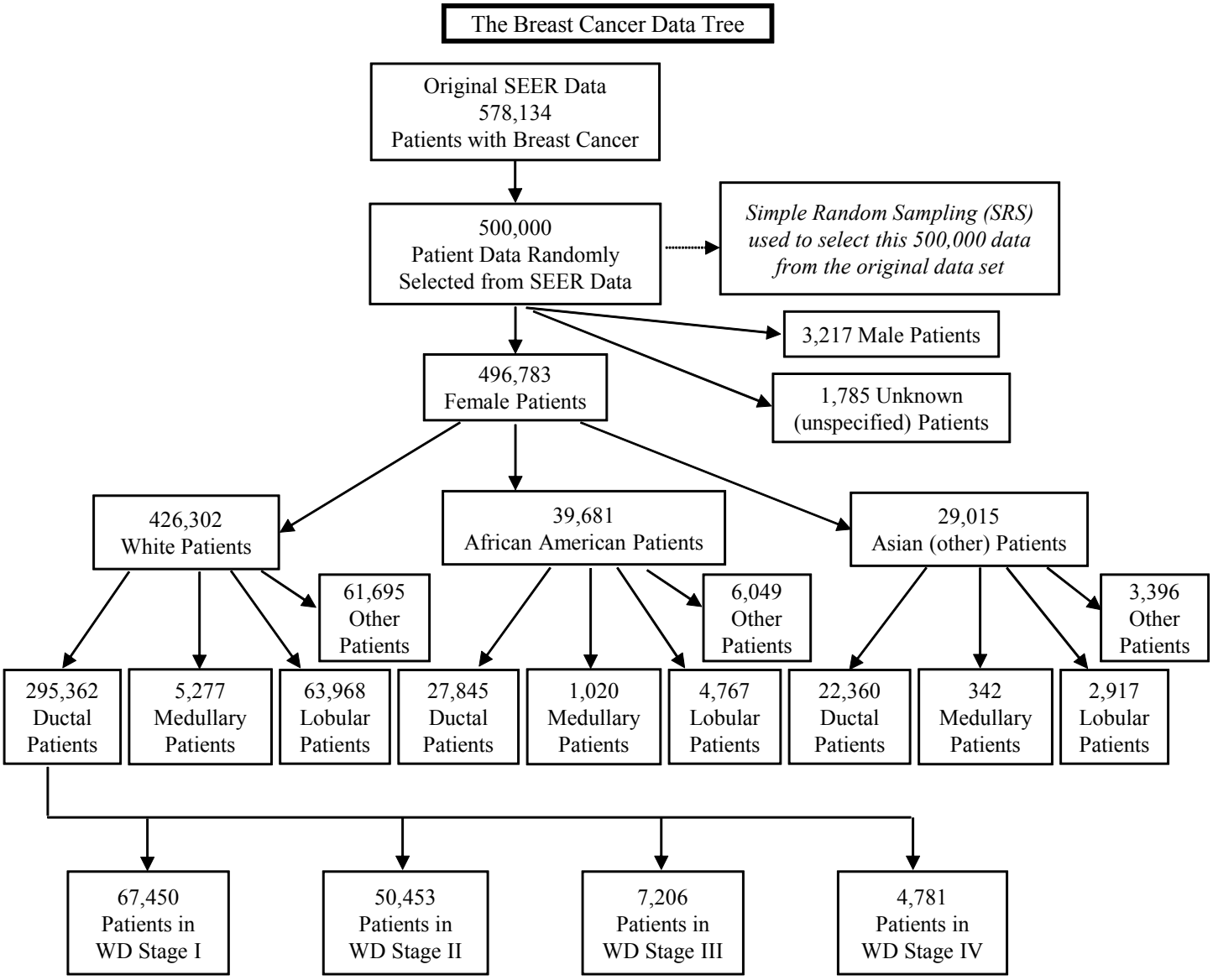
3. If $s \leq t$, then $N(s) \leq N(t)$.

If $s \prec t$, then $N(t) - N(s)$ is the number of events occurring during the interval $(s,t]$.

A Poisson process is a stochastic process in which events occur continuously and independently of one another. The Poisson process is a collection {N(t): t ≥ 0} of random variables, where N(t) is the number of events that have occurred up to time t (starting from time 0). The number of events between times a and b is given as N(b) − N(a) and has a Poisson distribution. Each realization of the process {N(t)} is a non-negative integer-valued step function that is non-decreasing.

For NHPP, the rate parameter may change over time. In this case, the generalized

Figure 1: Breast Cancer Data Tree Diagram
(WD stage I stands for White Ductal cancer patients in AJCC Stage I)



rate function is given as λ(t), thus, the expected number of events between time a and time b is:

$$\lambda_{a,b} = \int_a^b \lambda(t)dt. \quad (1)$$

Therefore, the number of arrivals in the time interval (a, b], given as N(b) − N(a), follows a Poisson distribution with associated parameter λ, a, b as:

$$P[(N(b)) - N(a)) = k] = \frac{e^{-\lambda_{a,b}}(\lambda_{a,b})^k}{k!}, \quad k = 0,1,... \quad (2)$$

A homogeneous Poisson process may be viewed as a special case when λ(t) = λ, a constant rate.

Figure 2: Breast Cancer Data Diagram White Ductal Stage I Patients

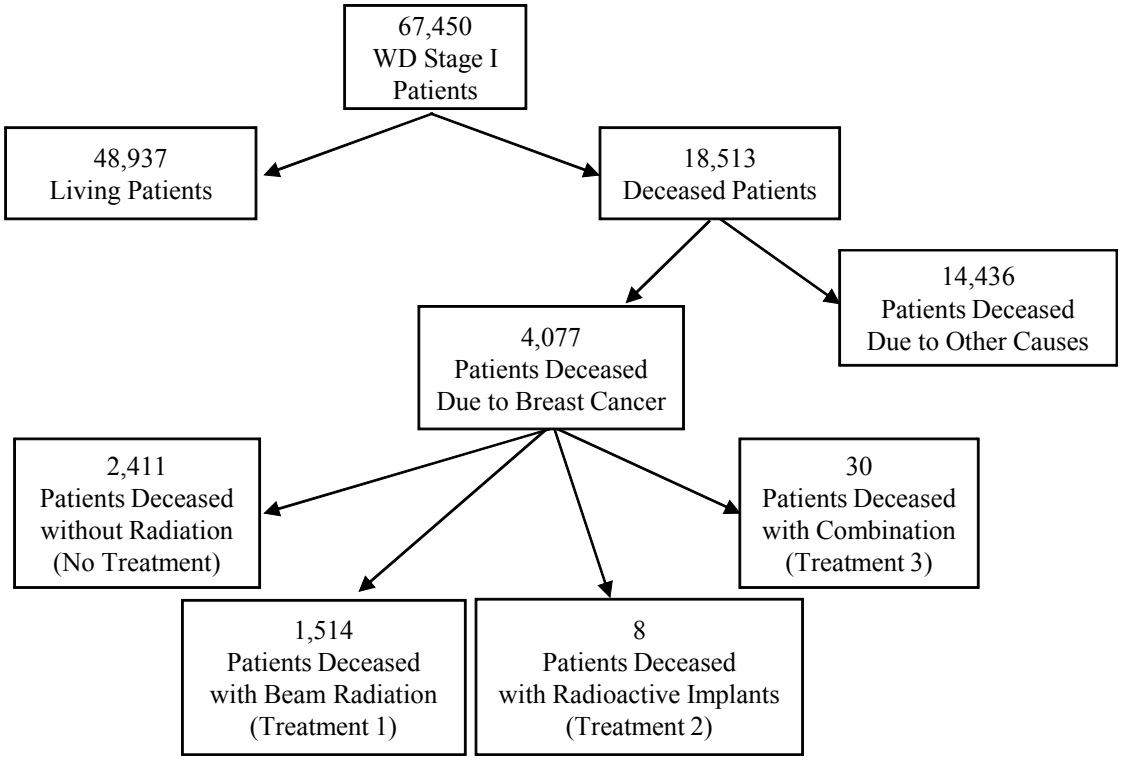
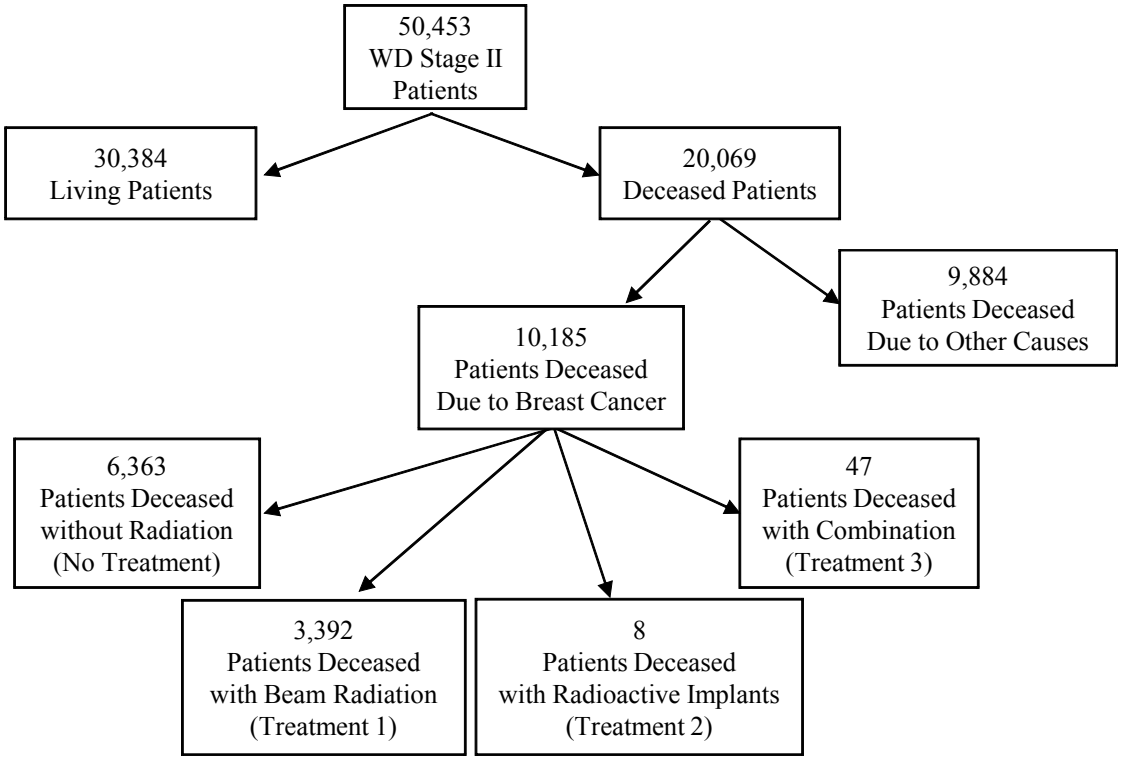


Figure 3: Breast Cancer Data Diagram White Ductal Stage II Patients



The mean value function $\lambda(t)$ of the process is:

$$\begin{aligned}\lambda(t) &= E(N(t)) \\ &= \int_0^t v(s)ds \\ &= \int_0^t \left(\frac{\beta}{\alpha}\right)\left(\frac{s}{\alpha}\right)^{\beta-1} ds \\ &= \left(\frac{t}{\alpha}\right)^{\beta}.\end{aligned} \quad (4)$$

It is known that, if the parameter beta is greater than one in survival analysis, then the failure time increases; this indicates a decrease in survival rate. If beta is less than one in the survival analysis, then the failure time decreases, meaning the survival rate increases. If beta equals one then the failure time is constant and the NHPP will become a homogenous Poisson process (HPP) (Rigdon & Basu, 2010).

The NHPP has the intensity function

$$v(t) = \left(\frac{\beta}{\alpha}\right)\left(\frac{t}{\alpha}\right)^{\beta-1}, \text{ for } \alpha \succ 0, \beta \succ 0, t \succ 0. \quad (3)$$

The unbiased estimator of beta is (Bain & Enelhardt, 1991):

$$\begin{aligned}\hat{\beta}_U &= \frac{n-1-\gamma}{n} \times \beta_{MLE} \\ &= \frac{n-1-\gamma}{\sum_{i=1}^{n} \log\left(\frac{t_n}{t_i}\right)}.\end{aligned} \quad (5)$$

where γ is an indicator function. If $\gamma = 1$ the system will be failure time truncated, meaning the system is restricted by a number of tails and testing will stop when that number of tails is reached. If $\gamma = 0$ then the system will be time truncated, which means the system is restricted by a final failure time and will stop when that time is reached.

The other parameter alpha can be calculated by equation 6, below.

$$\hat{\alpha} = \frac{t_n}{n^{\frac{1}{\hat{\beta}}}} \quad (6)$$

This study belongs to the first case; that is, the time of cases has been fixed. Patients were divided into four groups according to their cancer stage and, within each stage, it is known what kind of treatment the patient received, including if the patient did not receive any radiation treatment at all. Therefore, within each stage patients are divided into four groups with respect to treatment they received, namely, without treatment, treatment 1, 2 or 3. Treatment 1 refers to beam radiation, treatment 2 refers to radioactive implants and treatment 3 is a combination treatment. Few patients in the data source had treatments 2 or3, thus, those are the smallest groups.

Results

After calculating alpha and beta values for the NHPP for each treatment, results were compared and emerging patterns observed. Because the Caucasian race is the major population and ductal patients are the dominate type, this study focused on Caucasian ductal breast cancer patients. The estimation of the parameter is shown in Table 1.

Figure 4 shows the pattern for the key parameter beta. For example, β_{11} is 1.11 which means if a patient does not receive any treatment, the patient's condition will likely become worse because this indicates tumor growth which will lead to the progression of cancer. It may lead the patient to move from stage I to stage II or higher. Examining β_{31} and β_{32}, it is possible to determine whether a patient who receives treatment 3 in stage I will have a better result than a patient who receives the same treatment as a patient in stage II.

It was found that, for cases when beta are less than one, a decreased tumor size is indicated, meaning the treatment for breast cancer works. Results show that patients in early stages (for example, I and II) without treatment will experience increased tumor size and shorter time until death (see Table 1). Beam radiation

Figure 4: Evaluation Chain for NHPP

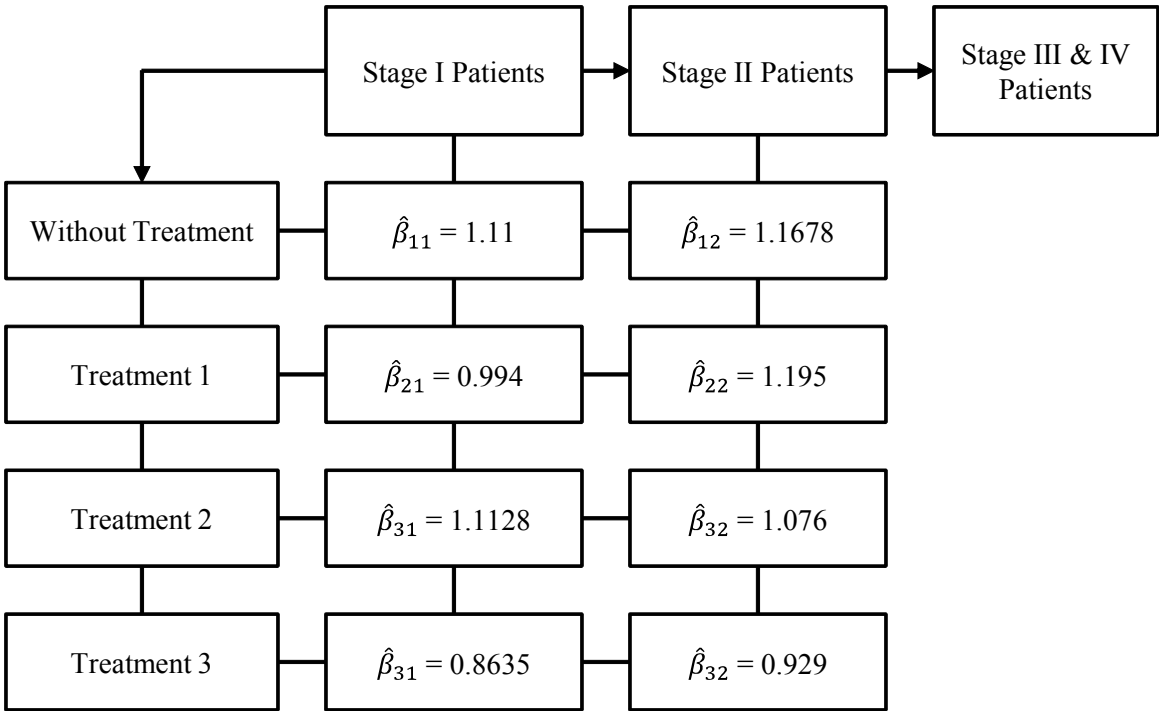


Table 1: Parameter Estimation for NHPP

		Stage I	Stage II
Alpha	Without Treatment	94.5112	113.8267
	With Treatment 1	56.17724	92.982
	With Treatment 2	76.03755	66.60
	With Treatment 3	33.8427	41.35
Beta	Without Treatment	1.110023	1.167756
	With Treatment 1	0.9943948	1.1195
	With Treatment 2	1.112772	1.076
	With Treatment 3	0.8635	0.929

(treatment 1) works for stage I but not for stage II. Radioactive implants (treatment 2) do not work well for either stage I or II. Treatment 3, a combination of the treatments, works well in stages I and II. (There is not enough data to conduct the NHPP for stages III and IV.) Intensity function plots are shown in Figures 5 - 12.

Figure 5 shows that, as the cumulative time of a patient increases, the intensity function also increases: this indicates, as expected, that tumor size is increasing and cancer is progressing. This result verifies the result obtained from parameter estimate β_{11}. Figure 6 shows that, as the cumulative time of a patient increases, the intensity function also decreases; this indicates, as expected, that the cancer will decrease with treatment 1 for stage I ductal Caucasian patients. This result leads to the same result obtained from parameter estimate β_{12}.

Figure 7 shows that, as the cumulative time of a patient increases, the intensity function also increases; this indicates, as expected, that the cancer progresses without treatment. This result verifies the result obtained from parameter estimate β_{13}. Figure 8 shows that, as the cumulative time of a patient increases, the intensity function decreases; this indicates that the cancer will improve with treatment 1 for stage 1 ductal Caucasian patients. This result leads to the same result obtained from the parameter estimate β_{14}.

Following a similar method, Figures 9, 10 and 11 show that, as the cumulative time of a patient increases, the intensity function also increases. This indicates that the cancer progresses without treatment or with treatment 1 or 2 for stage II ductal Caucasian patients. This result leads to the same result obtained from the parameter estimates β_{21}, β_{22} and β_{23}.

Figure 12 shows that, as the cumulative time of a patient increases, the intensity function decreases, this indicates - as expected - that the cancer will improve with treatment 3 for stage II ductal Caucasian patients. This result attests to the estimation obtained from parameter estimate β_{24} (see Table 1).

In summary, results indicate that, for Caucasian ductal breast cancer patients, it would be recommended to provide either a combination or a beam radiation treatment when they are in early stages I and II.

Conclusion

Based on breast cancer patients from the SEER database, adequate data exists to apply the NHPP analysis to Caucasian ductal cancer female patients in two early stages. Based on the results obtained from applying the proposed model, the following conclusions are put forth:

- With no treatment, the intensity function in stage I and stage II increases exponentially, implying that the tumor size of the patients increases at the same rate.
- With treatment 1 (beam radiation) in stage I the intensity function decreases, implying that the tumor size decreases. However, the same treatment in stage II shows the opposite result.
- With treatment 2 (radioactive implants) the intensity function in stage I increases and similar behavior is observed for the same treatment in stage II, this implies that the tumor size of the patients increases at the same rate.
- With treatment 3 (combination treatment) the intensity function in stages I and II decreases exponentially, this implies that the tumor size of the patients decreases at the same rate.

The study reported here is part of a larger, ongoing study. We will continue to obtain data and, eventually to construct a NHPP for each stage and each tumor size available for all treatments and compare the results. With more data and a broader range of patients and cancer stages, it will be possible to make suggestions for the particular treatment that will be best for patients with a particular tumor size. NHPP may also be applied to Bayesian survival analysis to compare and improve results.

Figure 5: Stage I Breast Cancer Intensity Function without Treatment

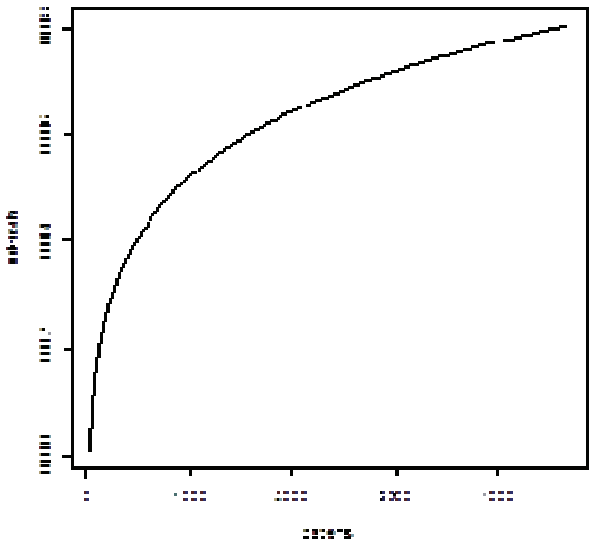


Figure 6: Stage I Breast Cancer Intensity Function with Treatment 1

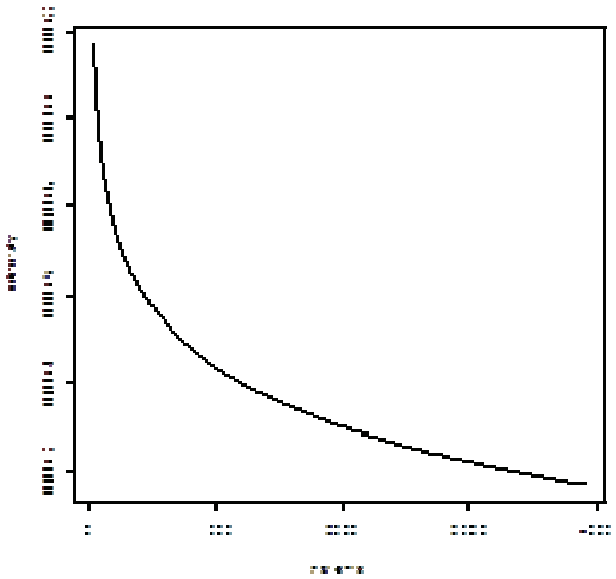


Figure 7: Stage I Breast Cancer Intensity Function with Treatment 2

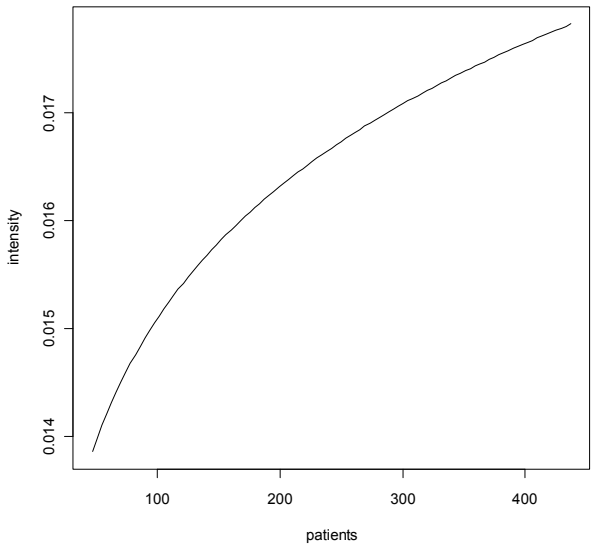


Figure 8: Stage I Breast Cancer Intensity Function with Treatment 3

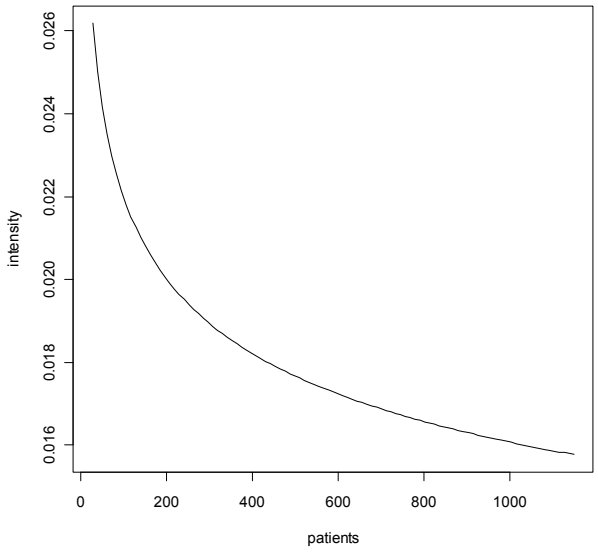


Figure 9: Stage II Breast Cancer Intensity Function without Treatment

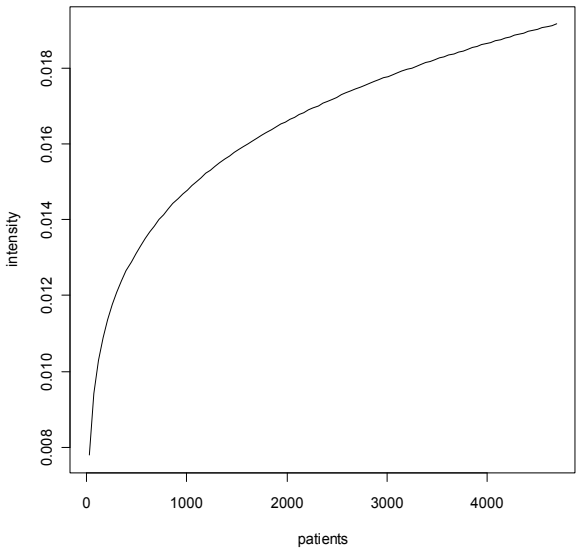


Figure 10: Stage II Breast Cancer Intensity Function with Treatment 1

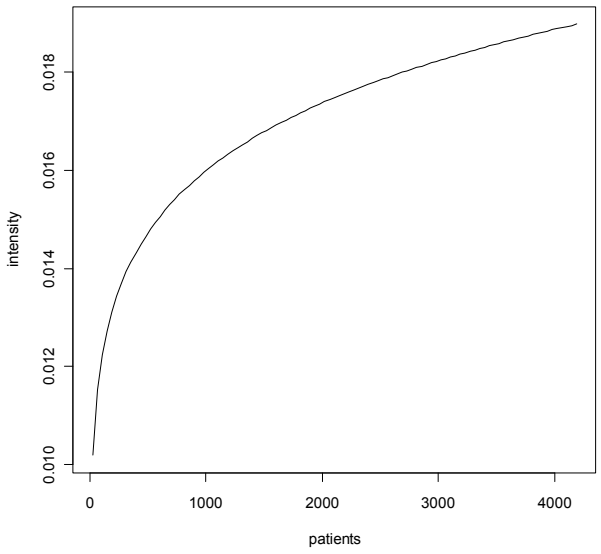


Figure 11: Stage II Breast Cancer Intensity Function with Treatment 2

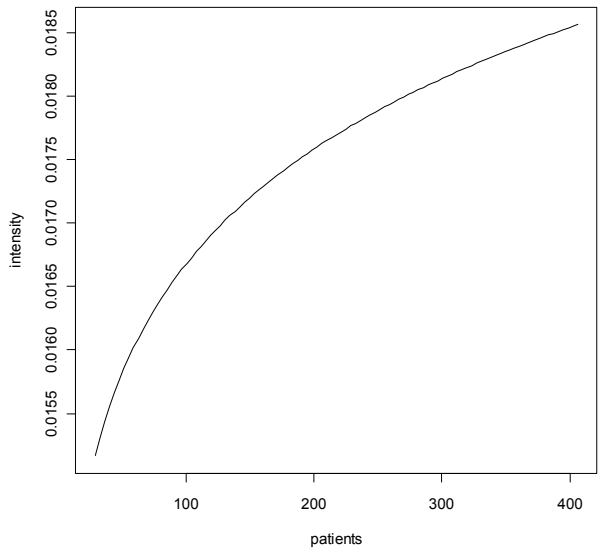
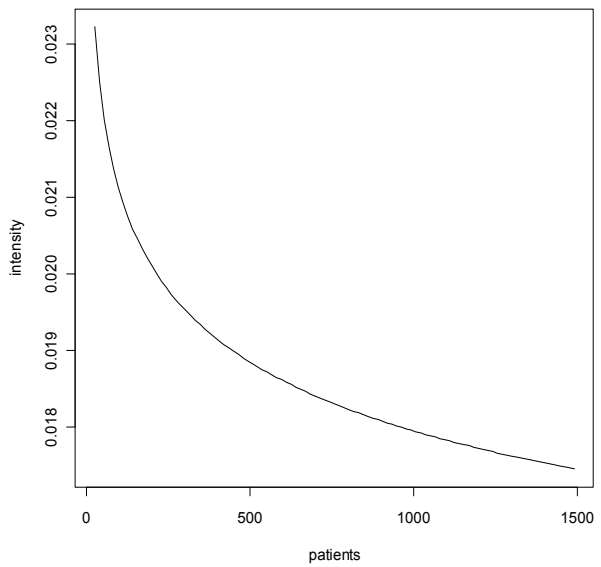


Figure 12: Stage II Breast Cancer Intensity Function with Treatment 3



References

Bain, L. J., & Enelhardt, M. (1991). *Statistical analysis of reliability and life testing models*, *2nd Ed*. New York, NY: Marcel Dekker.

Basawa, I. V., & Rao, B. L. S. P. (1980). *Statistical inference for stochastic processes*. London, England: Academic Press.

Billingsley, P. (1961). *Statistical inference for Markov processes*. Chicago, IL: University of Chicago Press.

Cox, D. R., & Lewis, P. A. W. (1966). The statistical analysis of series of events. *Annals of Mathematical Statistics*, *37*(*6*), 1852-1853.

Cox, D. R., & Isham, V. (1980). *Point processes*. London, England: Chapman & Hall.

Dharmadhikari, A. D., et al. (1989). Estimation of the scale parameter of a power law process using power law counts. *Annals of the Institute of Statistical Mathematics*, *41*(*1*), 139-148.

Duane, J. T. (1964). Learning curve approach to reliability monitoring. *IEEE Transactions on Aerospace*, *2*, 563-566.

Kingman, J. F. C. (1993). *Poisson processes*. Oxford, England: Oxford University Press.

Tsokos, C. P. (1997). Parameter estimation of power law process. In *Nonlinear Problems in Aviation and Aerospace*, 575-86. United States, CRC Press.

Rigdon, S. E., & Basu, A. P. (2000). *Statistical methods for the reliability of repairable systems*. New York, NY: Wiley.

Sariego, J. (2010). Breast cancer in the young patient. *The American Surgeon*, *76*(*12*), 1397-1400.

U. S. National Institutes of Health (2010). Access at: http://seer.cancer.gov.