# Treatment the Overdispersion in Survival Data Analysis (Suggested Model)

**Prof. Dr**

**Fatma A.Abdel-Elaty**
**Professor of Applied Statistics**
**Faculty of Commerce**
**Mansoura University**

**Dr**

**Ahmed Eltantawy Ismael**
**Lecturer of Oncology Medicine**
**Faculty of Medicine**
**Mansoura University**

**Researcher: Mennatallah Samir Ahmed Elhamady**

## Abstract

This research studies the over-dispersion on breast cancer patient data resulting from the presence of a latent variable and varying fluctuations around the mean. Both Fractional Poisson distribution and Standard Poisson distribution can be applied to estimate the number of positive axillary lymph nodes in the cases of the presence and absence of cancer stem cells at different time intervals for the size of the tumor. Additionally, The Relaxed Cure Rate model and the Standard Cure Rate model can be applied to estimate the survival probability and time of breast cancer patients when undergoing radiation therapy and when not undergoing radiation therapy in the presence or absence of cancer stem cells. Furthermore, survival rate models can be applied at different intervals for the number of positive axillary lymph nodes when radiation therapy was applied versus when it was not applied.

**Key Words**: Standard Poisson Distribution – Fractional Poisson Distribution – Relaxation Function – Relaxed Poisson Distribution – Stanadrad Cure Rate Model – Relaxed Cure Rate Model – Latent Cells – Latent Stem Cells.

## 1. Introduction

One of the leading causes of the emergence of the over-dispersion in data is the presence of bias in the resulting estimates and the difficulty in determining the standard error, thus resulting in inaccurate estimates of the parameters. This means an increase in the amount of variation from the expected value and a decrease in the correlation coefficient between response variables, followed by a decrease in reliability. The causes leading to the appearance of excessive data dispersion vary depending on the nature of the data and the statistical method used. The over-dispersion can be expressed by the dispersion equation, which represents the variance multiplied by the dispersion parameter. This parameter is then estimated, and control over the excess variance is achieved by determining an accurate value for this parameter that is suitable for the nature of the data.

It is known that Poisson distribution is used to express the probability of an event occurring during a specific time period (waiting time for the event), which is characterized by the property of equi-dispersion, meaning that the mean equals the variance. Usually, in Poisson distribution, the process of exponential distribution is applied to represent the variable of waiting time for the event. In the presence of the problem of over-dispersion, the variance of the distribution becomes greater than its expected value. Consequently, the distribution parameter becomes a random variable, meaning it is not independent of time and varies ($\lambda$). Also, the graphical representation of this data takes a non-exponential form, resulting in a non-homogeneous Poisson process. Therefore, the use of traditional Poisson processes is inappropriate for dealing with this type of data.

And then, many studies and research have led to the modifications of the traditional Poisson process by adding a new parameter to the distribution, taking into account the presence of over-dispersion known as the dispersion Parameter. Its value lies within a specified range, allowing control over data dispersion by choosing an appropriate value for this parameter within that range, reflecting the flexibility of the proposed distribution. Additionally, the use of a suitable distribution for the event waiting time, which considers the non-exponential shape of the data, should be employed. This distribution should possess mathematical properties that enable it to control the impact of over-dispersion on the event waiting time.

Laskin (2003) and Mauro et al. (2011) proposed the Fractional Poisson distribution as a generalization of the standard Poisson distribution by introducing an additional dispersion parameter to the distribution. They assumed that the observed waiting time follows the Exponential Mittag-Leffler distribution suggested by Leffler, M (1903), which is used to solve mathematical problems related to multidimensional functions and is a generalization of the exponential distribution function. Then, Rodrigues (2016) proposed the Relaxed Poisson distribution, where he introduced the Fractional Poisson distribution in the form of the Relaxed Mittag-Leffler function by Berberan-Santos (2005), which is characterized by its ability to change shape by controlling a parameter value within a certain range [0,1]. At the lower bound of the function, it takes the form of the geometric function, representing exponential decay when graphically represented. At the upper bound of this range, the function returns to the form of the exponential distribution.

The Relaxed Poisson Distribution is characterized by its ability to be represented in the space of complex numbers (which includes both imaginary and real numbers). However, representing some latent data with small weights in the space of imaginary numbers leads to bias in the resulting estimates.

B. Davies, Martin (1979), J. L. Brancik (1998), M. N. Berberan-Santos (2005) proposed an alternative representation on the space of positive real numbers for the Mittag-Leffler function. We can derive this integral representation for the Relaxed Poisson Distribution on the space of positive real numbers.

The traditional parametric and non-parametric survival functions assume that every individual in the study population is susceptible to the event, they do not represent surviving units with an infinite survival time.

Berkson and Gange (1952), Baog (1949), Yakovlev et al. (1996), Yakovlev (1994), and Tsodikov (1993) proposed models known as Mixture Cure Rate Models and Non-Mixture Cure Rate Models. These models establish a mathematical relationship representing a portion of the population that survives with an infinite survival time, and another portion representing the non-surviving population with a finite survival time.

The study aims to develop a model for analyzing survival data using the Non-Mixture Cure Rate Model for data related to the recurrence of breast cancer tumors in breast cancer patients. The goal is to study the likelihood of the tumor recurrence after radical mastectomy and radiation therapy, or in cases where the patient does not receive radiation therapy, due to the presence of dormant stem cancer cells, which represent the latent variable. The observed waiting time is the time until the event of the reappearance of the cancerous tumor, and the event under study is the recurrence of the cancerous tumor.

Radiation therapy is considered one of the treatment modalities used in breast cancer, where radiation therapy ions destroy the DNA of cancer cells. However, dormant cancer cells can resist targeted radiation therapy by halting their growth cycle for a period of up to years, then resuming growth and division to form a new cancerous tumor.

Cancer stem cells are considered the most dangerous type of cancer cells as they are characterized by self-renewal and rapid spread. They are also among the types of cancer cells that enter a dormant state, leading to the emergence of new cancerous tumor. Biological evidence of the presence of cancer stem cells increases the likelihood of the recurrence of the cancerous tumor.

Axillary lymph nodes are considered one of the most important prognostic factors in breast cancer. They are immune glands located under the arms and represent the closest area to the breast circumference, making them the most likely region for cancerous tumor metastasis outside the breast. Estimating the number of affected (positive) axillary lymph nodes is a crucial factor in determining the severity of the cancerous tumor and its ability to spread.

The number of positive axillary lymph nodes is affected by the tumor size, where the relationship between the number of positive axillary lymph nodes and tumor size is linear at some tumor size intervals and nonlinear at others. The variation in the relationship between the number of positive axillary lymph nodes and tumor size leads to data dispersion when estimating the number of positive axillary lymph nodes away from its arithmetic mean.

In this research, we will focus on estimating the number of positive axillary lymph nodes using both the Fractional Poisson distribution and the standard Poisson distribution.

So, the research aims to achieve the following points:

- Applying the Relaxed Non-Mixture Cure Rate Model, which represents the Non-Mixture Cure Rate Model on the form of the Mittag-Leffler function. This model is characterized by the flexibility of the Mittag-Leffler function between geometric and exponential shapes based on the dispersion parameter. It considers the over-dispersion in the data resulting from the presence of the latent variable, representing dormant cancer stem cells, to analyze survival in breast cancer patients after radical mastectomy, whether they receive radiation therapy or not.
- Conducting a comparison between the Relaxed Non-Mixture Cure Rate Model and the Non-Mixture Cure Rate Model to assess the capability and accuracy of the standard Cure Rate Model in comparison to the relaxed form of the Mittag-Leffler function for dealing with the over-dispersion resulting from the presence of the latent variable.
- Estimating the survival time for breast cancer patients after radical mastectomy until the occurrence of new cancerous tumor, considering whether the patient receives radiation therapy or not and whether biological evidence of the presence of cancer stem cells exists or not, and at different intervals of the number of positive axillary lymph nodes.
- Estimating the number of Positive Axillary Lymph Nodes using the Fractional Poisson distribution, and comparing it with the Standard Poisson distribution to assess the capability of the fractional form of the Poisson distribution in dealing with data on the number of positive axillary lymph nodes that exhibit significant variance from their arithmetic mean.
- Finding the integral image of the Relaxed Poisson distribution on the positive real number space to enable the model to deal with a wider range of data.

# 2. The Non-Mixture Cure Rate Models

## a. The Relaxed Non- Mixture Cure Rate Model

Rodrigues (2016) presented the survival function of the cure rate model in the form of the Mittag-Leffler function as follows:

$$S_{\phi}(t) = P[T > t] = A_M(S(t); \lambda, \phi) = E_{\phi}\big(-\lambda F(t)\big);\ \phi \in [0, 1] \tag{1}$$

As:

$\phi$ is the dispersion parameter,

$\lambda$ is the model rate,

t is the survival time,

$F(t)$ is the cumulative.

## b. The Standard Non- Mixture Cure Rate Model

The Standard Non-mixture Cure Rate model provides from the following relationship:

$$\mathrm{s(t)} = \exp\big(-\lambda \mathrm{F(t)}\big) \tag{2}$$

As:

$\lambda$ is the model rate,

t is the survival time,

$F(t)$ is the cumulative.

## c. Model Estimation

Both the Non-Mixture Cure Rate model and the Mixture Cure Rate model will be estimated, by the Maximum Likelihood function for the Non-Mixture Cure Rate model as follows:

$$L(\boldsymbol{\vartheta}; \boldsymbol{D}) = \prod_{i=1}^{n} \{\boldsymbol{f}(t;\ \boldsymbol{\vartheta})\}^{\delta_i} \{S(t;\ \boldsymbol{\vartheta})\}^{1-\delta_i} \tag{3}$$

As:

$\delta_i$ represents the indicator variable, taking the value 0 in the case of cure or no recurrence of cancer, and 1 in the case of cancer recurrence,

$\boldsymbol{\vartheta}$ represents the parameter vector,

$\boldsymbol{D}$ represents the variables vector,

$S(t;\ \boldsymbol{\vartheta})$ is the Non- Mixture Cure Rate model function,

$\boldsymbol{f}(t;\ \boldsymbol{\vartheta})$ is the density function of the Non- Mixture Cure Rate model.

And then the posterior distribution for the parameter vector $\boldsymbol{\vartheta}$ using Bayesian Inference method is given by:

$$\pi(\boldsymbol{\vartheta};\boldsymbol{D}) = \frac{[L(\boldsymbol{\vartheta};\boldsymbol{D})][\pi(\boldsymbol{\vartheta})]}{\int[L(\boldsymbol{\vartheta};\boldsymbol{D})][\pi(\boldsymbol{\vartheta})]} \tag{4}$$

As:

$\pi(\boldsymbol{\vartheta})$ represents the prior distribution for the vector parameter $\boldsymbol{\vartheta}$.

Both the Non-Mixture Cure Rate model and the Non-mixture Standard Cure Rate model are estimated using Markov Chain Monte Carlo (MCMC) methods applied to Bayesian inference, employing the Metropolis-Hastings algorithm to get the posterior distribution of model parameters.

### d. The Comparison Criteria between the Relaxed Non-Mixture Cure Rate Model and the Standard Non-Mixture Cure Rate Model:

The following comparison criteria are applied:

- The Akaike Information Criterion (AIC) value.
- The Bayesian Information Criterion (BIC) value.
- Calculating the bias resulting from each model.

## 3. The Standard Poisson Distribution and the Fractional Poisson Distribution

Applying both the Standard Poisson distribution and the Fractional Poisson distribution to estimate the number of positive axillary lymph nodes for breast cancer patients at different tumor sizes.

### a. Standard Poisson Distribution:

The function of the Standard Poisson distribution is as follows:

$$p(X = x;\ \lambda) = \frac{\lambda^x\ e^{-\lambda}}{x!}\ ;\quad \lambda = \exp(\boldsymbol{\beta} y_i^T) \tag{5}$$

As:

x is the discrete random variable representing the number of the positive axillary lymph nodes,

$\lambda$ is the distribution rate,

$\boldsymbol{\beta}$ is the parameter vector,

$y_i^T$ represents the covariates vector.

## b. Fractional Poisson Distribution:

The Following equation represents the Fractional Poisson distribution Laskin (2003), Mauro et al. (2011):

$$\Pr(x;\ \lambda, \phi) = \frac{\lambda^x}{x!} \sum_{k=0}^{\infty} \frac{(x+k)!}{k!} \frac{-\lambda^k}{\Gamma\phi(k+x)+1} \quad ;\quad x \geq 1,\ \ \phi\ \epsilon\ [0,1]$$

$$, \lambda = \exp(\boldsymbol{\beta} y_i^T) \tag{6}$$

As:

x is the discrete random variable representing the number of the positive axillary lymph nodes,

$\lambda$ is the distribution rate,

$\phi$ is the dispersion parameter,

$\boldsymbol{\beta}$ is the parameter vector,

$y_i^T$ represents the covariates vector.

## c. Estimation Method for the Standard Poisson distribution and the Fractional Poisson distribution:

Both the Fractional Poisson distribution and the Standard Poisson distribution are estimated using Markov Chain Monte Carlo (MCMC) methods, applied through the Bayesian approach, utilizing the Metropolis-Hastings algorithm to find the posterior distribution of the distribution's parameters.

## d. The Comparison Criteria between the Standard Poisson distribution and the Fractional Poisson distribution:

The following comparison criteria are applied:

- Calculating the Mean Square Error (MSE) for the distribution parameters.
- Calculating the Pearson Chi-Square test statistic to assess the ability of both the Fractional Poisson distribution and the Standard Poisson distribution to handle the over- dispersion.
-

## 4. The Relaxed Poisson Distribution:

Rodrigues (2016) displays the Relaxed Poisson distribution as follows:

$$P[M = m\,;\lambda,\phi] = \; p(m,\lambda,\phi) = \\ = \lambda^m\left[E^m_{\phi,\phi m+1}(-\lambda) - \lambda E_{\phi,\phi(m+1)}(-\lambda)\right];\;\; \phi \in [0,1] \tag{7}$$

As:

m is the Relaxed Poisson discrete random variable,

$\phi$ is the dispersion parameter,

$\lambda$ is the Relaxed Poisson distribution rate,

$E_{\phi,\phi(m+1)}(-\lambda)$ is the relaxed representation for the two parameters Mittag Leffler Function,

$E^m_{\phi,\phi m+1}(-\lambda)$ represents the relaxed representation for the three parameters Mittag Leffler function.

# 5. The Applied Study

The study was conducted on a total of 607 female breast cancer patients.

The duration of the survival after the radical mastectomy and receiving the radiation therapy or not receiving the radiation therapy ranged from 10 years from January 2012 to January 2022.

The study data was collected from the National Cancer Institute at Mansoura University.

The data is applied using R-Software version 4.2.1.

The research studies the estimation of the number of positive axillary lymph nodes of breast cancer patients, the practical application utilizes the Standard Poisson distribution and the Fractional Poisson distribution. This is based on the appearance of biological evidence in the study, such as the presence of cancer stem cells, and the absence of positive biological evidence in the study. This estimation is performed for three different tumor size intervals: less than 2 centimeters, 2 to 5 centimeters, and greater than 5 centimeters.

Applying the Improper Non-Mixture Standard Cure Rate model, the Relaxed Improper Cure Rate model, the Relaxed Cure Probability and the Standard Cure Probability model to estimate the survival probabilities in breast cancer patients after radical mastectomy. This estimation is performed in two scenarios: when radiation

therapy is applied and when radiation therapy is not applied. It is based on the appearance of biological evidence in the study for the presence of cancer stem cells, and the absence of biological evidence in the study for the presence of cancer stem cells. Both the application of the Standard Improper Non- Mixture Cure Rate model and the Relaxed Improper Non-Mixture Cure rate model, as well as the Standard Cure Probability and the Relaxed Cure probability, are applied to estimate the probabilities of survival for different numbers of positive axillary lymph nodes intervals: less than 3, 4 to 9, and greater than 9. This estimation is performed for both scenarios: when radiation therapy is applied after radical mastectomy and when radiation therapy is not applied after radical mastectomy.

# 6. Conclusions

- The flexibility of the relaxation function for the Non-mixture Cure Rate model between the exponential shape at the upper limit of the dispersion parameter and the geometric shape at the lower limit of the dispersion parameter, has made the function take into account the weight of sudden changes in the data in different directions around the mean. This has resulted in lower AIC and BIC criterion values and lower bias values compared to those resulting from the Non-Mixture Standard Cure Rate model.
- The value of the survival probability resulting from both the Relaxed Non-Mixture Cure Rate model and the Standard Non-Mixture Cure Rate model for breast cancer patients when there is biological evidence of the presence of cancer stem cells in the case of the patient receiving radiotherapy equals to 0.47 (applying the Relaxed Non- Mixture Cure Rate model) or the case of the patient not receiving radiotherapy equals to 0.42 is less than the probability of survival resulting from the absence of biological evidence of the presence of cancer stem cells with 0.54 and 0.5 respectively, as cancer stem cells are among the most dangerous types of cancer cells due to their rapid self-renewal and spread. Moreover, the likelihood of them entering a dormant state is high, which increases the resistance of cancer cells to targeted radiotherapy treatment.
- The value of the survival probability resulting from both the Relaxed Non-Mixture Cure Rate model and the Standard Non-Mixture Cure rate model for breast cancer patients when radiotherapy is applied, in the case of biological evidence of the presence of cancer stem cells or in the absence of biological evidence of the presence of cancer stem cells, is higher than the probability of survival resulting from not receiving radiotherapy. This is due to the effectiveness of radiotherapy in inducing cancer cell death through programmed cell death (apoptosis) after the destruction of their DNA, halting their growth and therefore proliferation, reducing the likelihood of cancer recurrence.
- The value of the survival Probability from both the Relaxed Non-Mixture Cure Rate model and the Standard Non-Mixture Cure Rate model for breast

cancer patients decreases with an increase in the number of positive axillary lymph nodes, whether the patient receives radiotherapy equals to 0.55 (in the interval less than 3), equals to 0.5 (in the interval from 4 to 9), and equals to 0.44 (in the interval greater than 9), greater than the survival probability in case of not applying the radio-therapy treatment, as the Relaxed Non-Mixture Cure Rate equals to 0.51 (in the interval less than 3), equals to 0.43 (in the interval from 4 to 9), and the survival probability equals to 0.37 (in the intervals greater than 9). An increase in the number of positive axillary lymph nodes indicates the ability of cancer cells to spread and divide outside the breast area, where cancer cells with the ability to spread are considered more dangerous.

- The survival time for breast cancer patients increases when the patient receives radiotherapy, whether cancer stem cells are present or not, compared to the scenario where the patient does not receive radiotherapy, regardless of the presence of cancer stem cells.
- The patient's survival time decreases for breast cancer patients when biological evidence of the presence of dormant cancer cells appears as it is estimated for 5 years in case of applying the radiotherapy treatment in the presence of cancer stem cells, less than the survival time estimated for 4 years in case of not applying the radiotherapy treatment in the presence of cancer stem cells, and the survival time is estimated for 6 years for the case of applying the radiotherapy treatment in case of applying the radiotherapy treatment in the absence of cancer stem cells which is bigger than the survival time that equals to 5 years in case of not applying the radiotherapy in case of absence of the cancer stem cells.
- The survival time decreases with an increase in the number of positive axillary lymph nodes.
- The results indicated that the Over- dispersion resulting from the Fractional Poisson distribution is lower than the Over- dispersion resulting from the Standard Poisson distribution when estimating the number of positive axillary lymph nodes for breast cancer patients using the Pearson's chi-square test as the values of Pearson Chi-Square test in case of the Fractional Poisson distribution is near to 1 which is the case of Equi-dispersion. Additionally, the mean squared error values for the parameters of the Fractional Poisson distribution are lower, given the flexibility of the distribution between exponential and geometric shapes at different dispersion parameter values, and the distribution's ability to represent non-exponential shapes of the data, making the Fractional Poisson distribution more optimal in representing data with excess dispersion.
- The average number of positive axillary lymph nodes increases with the size of the cancerous tumor. An increase in the size of the cancerous tumor signifies an increased likelihood of the number of cancer cells capable of spreading beyond the original breast tumor area.

- The average number of positive axillary lymph nodes increases when there is biological evidence of the presence of cancer stem cells compared to the average number of positive axillary lymph nodes when there is no biological evidence of the presence of cancer stem cells, as the number of estimated positive axillary lymph nodes resulting from the Fractional Poisson distribution in case of the presence of cancer stem cells, equals 4 (in the tumor size interval less than 2 cm), equals 5 (in the tumor size interval from 2 cm to 5 cm), and equals to 8 (in the tumor size interval greater than 5 cm) which is greater than the number of positive axillary lymph nodes in case of absence of the cancer stem cells resulting from the Fractional Poisson distribution, equals to 2 (in the tumor size interval less than 2 cm), equals to 3 (in the tumor size interval from 2 cm to 5 cm), and equals to 5 (in the tumor size interval greater than 5 cm).
- Cancer stem cells are characterized by their ability to self-renew and rapidly spread, increasing the likelihood of their migration to areas outside the original site of the cancerous tumor.

# 7. Recommendations

- Applying the Fractional Poisson distribution to breast cancer data with the phenomenon of Over- dispersion to estimate the number of positive axillary lymph nodes.
- Applying the Relaxed Non-Mixture Cure rate model to estimate and study the survival probability for breast cancer patients in the presence or absence of cancer stem cells and when applying radiotherapy or not, distinguished by its mathematical characteristics leading to statistically more significant results.
- Applying the radiotherapy treatment to breast cancer patients when there is biomarkers evidence of the presence of cancer stem cells, especially in cases of high positive axillary lymph node counts, particularly those exceeding nine affected axillary lymph nodes.

# 8. References


1- Alexander l. Saichev (1997). Fractional Kinetic equations: solutions and applications. American Institute of Physics. S1054-1500(97).

2- Alok Kumar Dwivedi, Sada Nand Dwivedi, Suryanarayana Deo, Rakesh Shukla, Elizabeth Kopras (2010). Statistical models for predicting number of involved nodes in breast cancer patients. Health Journal. Vol. 2, No. 7, 641-651.

3- Al Omari Mohammed Ahmed, Noor Akma Ibrahim, Mohd Bakri Adam, and Jayanthi Arasan (2012). Bayesian survival and hazard estimate for Weibull censored time distribution. Journal of Applied Sciences. 12, 1313-1317.

4- Amir Ahmad (2013). Pathways to breast cancer recurrence. ISRN Oncology Journal. Vol. 2013.

5- Anaid Anna Kasangain, et. al (2017). The prognostic role of tumor size in early breast cancer in the era of Molecular Biology. Plos One Journal. 12 (12).

6- Alexandre Lafourcade, Mathilde His, Laura Baglietto, Marie Christine Boutron Ruault, Laure Dossus, and Virginie Rondeau (2018). Factors associated with breast cancer recurrences or mortality and dynamic prediction of death history of cancer recurrences: the French E3N cohort. BMC Cancer Journal. 18:171.

7- Bernard Fisher, Madeline Bauer, Lawrance Wickerham, Carol K. Redmond, and Edwin R. Fisher (1983). Relation of number of positive axillary nodes to the prognosis with primary breast cancer. Cancer Journal. 52: 1551-1557.

8- Charmine B. Dean and Erin R. Lundy (2016). Overdispersion. Wiley StatsRef: Statistics Reference Online.

9- Cordell Gilreath, Marjan Boerma, Zhiqiang Qin, M. Keith Hudson, and Shanzhi Wang (2021).The hypoxic microenvironment of breast cancer cells promotes resistance in radiation therapy. Frontiers in Oncology Journal. Vol. 10.

10- Durga H. Kutal, and Lianfen Qian (2018). A non- mixture cure model for right censored data with frechet distribution. Stats Journal. 1, 176-188.

11- E.C. Grigoletto, E.C. Olivera, and F.F. Camargo (2019). Integral representation of Mittag-Leffler function on positive real axis. Tendencias em Mathematica Aplicada e Computacional Journal. Vol. 20, No. 2, 217-228.

12- Erica Pranzini, Giovanni, and Maria Letizia Taddei (2022). Metabolic features of tumor dormancy: possible therapeutic strategies. Cancers Journal. 14, 547.

13- E. Oricchio, J. Hulsken, K. De Visser, T. Petrova, M. De Palma (2022). The role of eosinophils in breast cancer. Faculte des sciences de la vie.

14- Fabiana Tonellotto, Anke Bergmann, Karn de Souza Abrahao, Suzana Sales de Aguiar, Marcelo A Bello, and Luiz Claudio Santos Thuler (2019). Impact of Number of positive lymph nodes and lymph node ratio on survival of women with node positive breast cancer. Breast Health Journal. Vol. 15, No. 2, 76-84.

15- Harry Bartelink, et. al (2001). Recurrence rates after treatment of breast cancer with standard radiotherapy with or without additional radiation. The New England Journal of Medicine. Vol. 345, No. 19.

16- Hui Zahang, Stanely B. Pounds and Li Tang (2013). Statistical Methods for Over-dispersion in mRNA seq Count Data (2013). The Open Bioinformatics Journal. 7, 34-40.

17- Hidetka Shima, Akimitsu Yamada, Takashi Ishikawa, Itaru Endo (2017). Are breast cancer stem cells the key to resolving clinical issues in breast cancer therapy? Gland Surgery Journal. Vol. 6, No. 1.

18- Hiroko Endo, Masahiro Inoue (2018). Dormancy in cancer. Wiley Cancer Science Journal. 110, 474-480.

19- Ibtisam Lale Atahan, Ferah Yildiz, Gokhan Ozyigit, Sait Sari, Murat Gurkaynak, Ugur Selek, and Mutlu Hayran (2009). Percent positive axillary lymph Node metastasis predicts survival in patients with non-metastatic breast cancer. Acta Oncologica Journal. No. 47, 232-238.

20- Jia-Long Wu, Hsin- Shun Tseng, Li-Heng Yang, Hwa-Koon Wu, Shou-Jen Kuo, Shou- Tung Chen, and Dar-Ren Chen (2014). Prediction of axillary lymph node metastases in breast cancer patients based on pathologic information of the primary tumor. No. 20, 577-581.

21- Josemar Rodrigues, Gauss M. Corderio, Vicente G. Cancho, and N. Balakrishnan (2016). Relaxed cure rate model. Biometrical Journal. 58, 397-415.

22- Julia Tutzauer, Martin Sjostrom, Erik Holmberg, Per Karlsoon, Fredrika Killander, L. M, Fredrik Leep Lundeberg, Per Malmstrom, Emma Nimeus, Marten Ferno, and Annika Jogi (2021). Breast cancer hypoxia in relation to prognosis and benefit from radiotherapy after breast conserving surgery in a large randomized trial with long-term follow up. British Journal of Cancer. 126: 1145-1156.

23- Komal Goel, Manoj Kumar Varshney, Gurprit Grover, and Seema Pant (2022). Om the estimation of cure fraction under power compertz distribution under Bayesian approach. Stochastic Modeling and Application Journal. Vol. 26, No. 1.

24- Melvin J. Silverstein, Kristin A. Skinner, and Thomas J. Lomis (2001). Predicting axillary nodal positivity in 2282 patients with breast carcinoma. World Journal of Surgery. 25, 767-772.

25- Mario N. Berberan- Santos (2005). Properties of the Mittag-Leffler relaxation function. Journal of Mathematical Chemistry. Vol. 38, No. 4.

26- Manish Kumar Goel, Paedeep Khanna, Jugal Kishore (2010). Understanding survival analysis: Kaplan- Meier estimate. International Journal of Ayurveda Research. Vol 1,Issue 4.

27- Malgorzata Banys, Andreas D Hartkopf, Natalia Krawczyk, and Tatjana Kaiser (2012). Dormancy in breast cancer. Breast Cancer: Targets and Therapy Journal. 4, 183-191.

28- Mitra Rahimzadeh, Behrooz Kavehie, Mohammad Reza Zali (2014). Cure models in analyzing long- term survivors. Transl Gastriotest Cancer Journal. 3(4), 149-154.

29- Nick Laskin (2003). Fractional Poisson process. Elsevier Journal. Vol. 2003, No. 8, 201-213.

30- Paul C. Lambert, John R. Thompson (2006). Estimating and modeling the cure fraction in population- based cancer survival analysis. Biostatistics Journal. 8, 3, 576-594.

31- Romano Demicheli, Antonello Abbattista, Rosalba Miceli, Pinuccia Valagussa, and Gianni Bonadonnaa (1996). Time distribution of the recurrence risk for breast cancer patients under- going mastectomy: further support about the concept of tumor dormancy. Breast Cancer Research and Treatment Journal. 41: 177-185.

32- Siddhartha Chib and Edward Greenberg (1995). Understanding the Mteropolis- Hastings Algorithm. The American Statistician Journal. Vol. 49, No. 4.

33- Sibel Kahraman Cetintas, Meral Kurt, Lutfi Ozkan, Kayihan Engin, Sehsuvar Gokgoz, and Ismet Tasdelen (2006). Factors influencing axillary node metastasis in breast cancer. Tumori Journal. 92: 416-422.

34- Sungduk Kim, Yingmei Xi, and Ming-Hui Chen (2009). A new latent cure rate marker model for survival data. The Annals of Applied Statistics. Vol. 3, No. 3, 1124-1146.

35- Shima Younespour, Elham Maraghi, Amal Saki Malehi, Maedeh Raissizdeh, Mohammad Seghatolesmi, and Mehran Hosseinzadeh (2020). Evaluating Related Factors to the Number of Involved Lymph Nodes in Patients with Breast Cancer using Zero- Inflated Negative Binomial Regression Model. Biostat Epidemiol Journal. Vol. 6, No. 4, 259-266.

36- Shideh Rafati, Mohammad Reza Baneshi, Laleh Hassani, Abbas Bahrampour (2021). Survival analysis in the presence of a cure fraction using data of analysis patients: A Bayesian Approach. Research Square Journal.

37- Stefan Steurer, et. al. (2021). P63 expression in human tumors and normal tissues: a tissue microarray study on 10,200 tumors. Biomarker Research Journal. Vol. 9, No. 7.

38- Tobias M. Weissenbacher, et. al. (2010). Multicemtric and multifocal versus unifocal breast cancer: is tumor-node metastasis classification justified? Breast Cancer Res Treat.122: 27-34.

39- Ting Liu, Jun Lv, and Yutao Qin (2017). Standarized tumor volume: an independent prognostic factor in advanced nasopharyngeal carcinoma. Ontotarget Journal. Vol. 8, No. 41.

40- Valerie W. Rush, John Growley, Dorothy J. Griroux, Peter Goldstraw, Jung-Gi Im, Masahiro Tsuboi, Ryosuke Tsuchiya, and Johan Vansteenkiste (2007). The IASLC lung cancer staging project: proposals for the revision of the N descriptors in the forthcoming seventh edition of the TNM classification for lung cancer. Journal of Thoracic Oncology. Vol. 2, No. 2.

41- Vicki Plaks, Niwen Kong, and Zena Werb (2015). The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? Elsevier Inc Journal. Vol. 16, No. 3, 225-238.

42- Veronica Gatti, Lucilla Bongiorno-Borbne, Glaudia Fierro, Margherita Annicchiarico- Petuzzelli, Gerry Melino, and Angelo Peschiaroli (2019). P36 at the crossroads between stemness and metastasis in breast cancer. International Journal of Molecular Sciences. 20, 2683.

43- Vicente G. Cancho, Gauss M. Cordeiro, Gladys Barriga, Edwin M. M. Ortega, and Michael W. Kattan (2021). The destructive cure rate regression in cancer prognosis and prediction. Applied Mathematics and Information Sciences Journal. 2, 199-206.

44- Xiao-Lei Gao, Mei Zhang, Ya-Ling Tang, Zin-Hua Liang (2017). Cancer cell dormancy: mechanisms and implications of cancer recurrence and metastasis. Onto Targets and Therapy. 10, 5219-5228.

45- Wolfram Malter, Martin Hellmich, Mayhar Badian, Verena Kiran, PeterMallman, and Stefan Kramer (2018). Factors predictive of sentinel lymph node involvement in primary breast cancer. Anticancer Research Journal. No. 38, 3657-3662.

46- Wafa Boulefour, Elise Rowinski, Safa Louati, Sandrine Sotton, Anne- Sophi Wozny, Pablo Moreno-Acosta, Benoite Mery, Claire Rodriguez-Lafrasse, and Nicolas Magne (2021). A review of the role of hypoxia in radiotherapy in cancer therapy. Medical Science Journal. Vol. 27.

47- Young Ju Song, Sun Hyoung Shin, Jin Seong Cho, Min Ho Park, Jung Han Yoon, and Young Jong Jegal (2011). The role of lymph-vascular invasion as a prognostic factor in patients with lymph node- positive operable invasive breast cancer. Journal of Breast Cancer. 14(3): 198-203.

48- Yuanxin Zhang, Ji Li, Yuan Fan, Xiaomin Li, Juanjuan Qiu, Mou Zhu, and Hongjiang (2019). Risk factors for axillary lymph node metastases in clinical stage T1- 2N0M0 breast cancer patients. Medicine Journal. 98:40.

49- Yangfan Liu, Miao Yang, Jingjing Luo, and Hongmei Zhou (2020). Radiotherapy targeting cancer stem cells "awakens" them to induce tumor relapse and metastasis in oral cancer. International Journal of Oral Science. 12:19.

50- Yin Liu, Min He, Wen-Jia Zuo, Shuang Hao, Zhong- Hua Wang, and Zhi Ming Shao (2021). Tumor size still impacts prognosis in breast cancer with extensive nodal involvement. Frontiers in Oncology Journal. Vo. 11.