WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Stochastic Modeling For Nanosecond Pulsed Electric Fields Inhibit Breast Cancer Development

***Dr.N.Umamaheshwari* [1] *and E.Thatchayeni* [2]**

[1] Assistant proffessor, Department of Mathematics, Shrimati Indira Gandhi College, Affliated to Bharathidasan University, Trichy, Tamilnadu, India.

[2] M.Phil Scholar, Department of Mathematics, Shrimati Indira Gandhi College, Affliated to Bharathidasan University, Trichy, Tamilnadu, India.

ABSTRACT

Nanosecond Pulsed Electric Fields (nsPEFs) is a novel non-thermal approach to induce cell apoptosis. In this study, we established breast cancer animal model with MCF-7 cell line on Balb/c nude mice. 30 tumor were exposed to 720 pulse of 100ns at 30kV/cm. Treated tumor was inhibited by 79% 2 week after pulse. A 3.0T clinical magnetic resonance imaging (MRI) system with the self-made surface coil was employed to observe morphological changes of tumors. Pulsed tumors showed histological markers of apoptosis and decreased blood vessel density. To evaluate the might-be adverse effects ofnsPEFs in healthy tissues, normal skin was treated in exactly the same way as tumors, and pulsed skin showed no permanent damages. The results suggest nsPEFs is able to inhibit human breast cancer development and suppress tumor blood vessel growth, indicating nsPEFs may serve as a novel therapy for breast cancer in the future.

Keywords: Lograthamic transformation, geometric mean analysis, ex-Gaussian distribution, log-normal distribution.

INTRODUCTION

Breast cancer is one of the most threatening malignant tumors among women, the incidence of which is rising year by year. Despite of early screening and improvement in breast cancer management that have increased the 5- year survive rate, the requirement for novel and more efficient therapy for breast cancer is still quite urgent.

Breast cancer now ranks first as a cause of death of all cancers among women between 20 to 59 years old [1]. There are several non-thermal therapies applied in clinics or being tested in laboratory. Electrochemotherapy and electro-gene therapy both use electroporation to introduce chemotherapeutic drugs or toxin genes into tumor cells with electric pulses of 100us duration [2][3].

Numerous significance tests assume data are normally distributed such as t-tests, chi-square tests and f- tests. This is often reasonable, as many real-world measurments/observations follow a normal distribution; however, there are several situations in which the normal distribution assumption is not appropriate (eg., immunologic and reaction time data). The normal distribution, sometimes called a Gaussian distribution, is characterised by symmetric, Bell-like shape.

There are few broad, well-established guidelines for choosing the most appropriate transformation. For instance, a logarithmic transformation is recommended for positively skewed data, while negatively skewed data is often transformed using the square root function. The ex-Gaussian distribution is another method, often used for negative, positively skewed data. This distribution is defined by three parameters; and an ex-Gaussian analysis involves estimation of these parameters usually by either method likelihood estimation (Heathcote, 1996). This paper will first introduce basic statistical methodology that is essential in understanding the transformation of data and the ex-Gaussian distribution then the Box-Cox family of transformations and methods of estimating the ex-Gaussian distribution will be introduced, and will conclude with conclude with conclude with a discussion regarding the usefulness of both methods.

STASTICAL BACKGROUND

Symmetric data are often best described by a measure of centre (e.g., mean) and by a measure of spread (e.g., standard deviation). skewed data, on the other hand, benefit from describing the direction (positive or negative) and magnitude of skew. Describing the skewness of a data set is a concept that is often overlooked. In theoretical terms, a distribution's skew iw defined as

$$\gamma_1 = \frac{\mu_3}{\sigma_3}$$

Where σ is the standard deviation and μ_3 is the third moment about the mean (variance is the second moment about the mean, i.e., $\mu_2 = \sigma^2$ and $\gamma_2 = \frac{\mu_4}{\sigma_4}$ is known as *kurtosis*). Kurtosis is a measure of the thickness of a distribution's tails and thus affects how p-values are computed. However, kurtosis is not as important as skewness when performing normal distribution-based significance tests such as t-tests. T-tests are protective against excess kurtosis (i.e., heavier tails) because probabilities compute in the tails of the t-distribution are larger than those from the normal distribution making the analysis more conservative.

Values of γ_1 can be either negative or positive which correspond respectively to negative or positive skewness (sometimes referred to as left and right skewed respectively). Skewness is different for each probability distribution and can be computed from a distribution's probability density function. To get an understanding about probability density function. To get an understanding about how to analyse a data set, we will next discuss the properties of four probability distribution functions, namely the normal, log-normal, exponential, ex-Gaussian distribution. This is an abbreviated description and the authors recommend Wackerly, Mendenhall and Scheaffer (2007) for a more complete discussion of mathematical statistics.

The Normal Distribution

The first probability distribution that will be discussed is the normal distribution, the backbone of statistical analysis. The normal distribution characterised by a bell-like shape with a probability density function (pdf) that can be written as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)}{\sigma^2}}, \ -\infty < x < \infty$$

Where the mean μ and σ^2 uniquely defined distribution of the random variable X. A probability distribution ids often written as a pdf because this function is used to graph the probability distribution as a smooth curve for continuous data and useful information and skewness are computed form this function.

The normal distribution has no skew(i.e., $\gamma_1 = 0$), so it is a very poor choice to model positively skewed data. On the other hand, non-normal data can oft6en be transformed to a distribution that is reasonable normal. As mentioned above, transforming a non-normal set of data t a distribution that is reasonably normal can be very useful as many significance tests rely on assumption $\overline{X}$ is normally (which is certainly true when data has been sample from a normal distribution). These test include various forms t-test and analysis of variance (ANOVA). There are even several non-parametric test that delay on a relaxed normality assumption (man-whitny, Kruskal-wallis and chi-square space test).

The log-normal distribution

A variant of the normal distribution is the log-normal distribution. In statistical jargon, a random variable W with parameters μ and σ has a log-normal distribution if the random variable $U = \ln(W)$ has a normal distribution, where ***ln*** is the logarithm with base ***e***, i.e., **log₁e**and sometimes called the *natural logarithm*. In broad terms, a log-normal distribution is distribution that is normal when log-transformed. The pdf of the log-normal distribution can be written as

$$f(w) = \frac{1}{w\sigma\sqrt{2\pi}} e^{\frac{(\ln(w)-\mu)}{\sigma^2}}, \qquad 0 < w < \infty$$

Where μ and σ are the mean and standard deviation of the normal random variable U respectively . The mean and variance of the log-normal random variable W are

$\mu_w = e^{\mu+\frac{\sigma^2}{2}}$ and $\sigma_w^2 = \mu_w^2(e^{\sigma^2} - 1)$

Unlike the normal distribution, the log-normal distribution is positively skewed with $\gamma_1 = (e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}$, a value that is always positive when $\sigma > 0$. in practice , positively skewed data are often transformed using logarithms in an effort to make the data more reasonably normally distributed due to the relationship between the normal and log-normal distributions.

The Ex-Gaussian Distrbution

The Ex-Gaussian distribution is aptly named as it is the sum of two independent probability distributions, the exponential and the normal. The exponential distributions, the exponential and normal. The exponential distribution is a positively skewed distribution whose pdf for a random variable Y is

$$f(y) = \frac{1}{\lambda} e^{\frac{y}{\lambda}} \quad 0 < y < \infty$$

Where the mean λ uniquely defines this distribution. The variance and skewness of the exponential distribution are λ^2 and 2 respectively.

If X and Y represent independent normal and exponential random variables (as defined above), then Z=X+Y is an Ex-Gaussian random variable with mean and variance

$$\mu_z = \mu + \lambda \text{ and } \sigma_z^2 = \sigma^2 + \lambda^2$$

The skewness parameter for $\dot{z}$ is

$$\gamma_1 = \frac{2\lambda^3}{(\sigma^2 + \lambda^2)^{\frac{3}{2}}}$$

Which can be shown to never exceed 2, the skewness of the exponential distribution. The pdf of the Ex-Gaussian random variable Z is

$$f(z) = \frac{1}{\lambda}\phi\left(\frac{z-\mu}{\sigma} - \frac{\sigma}{\lambda}\right) e^{\left\{\frac{\sigma^2}{2\lambda^2} - \frac{z-\mu}{\lambda}\right\}}, \qquad -\infty < x < \infty$$

Where $\phi(.)$ is the cumulative distribution function of the standard normal distribution. The pdf of the Ex-Gaussian distribution can be viewed as a weighted average of the shifted exponential distribution, where

$$\frac{1}{\lambda} e^{\frac{z-\mu}{\lambda}}$$

Box-Cox transformation and Estimating the Ex-Gaussian Distribution

The previous section discussed many important statistical concept for dealing with positively skewed data. In this section, we will discuss method for dealing with a is an exponential distribution shifted by the normal mean parameter μ and weight

$$\phi\left(\frac{z-\mu}{\sigma}\right)e^{\frac{\sigma^2}{2\lambda^2}}$$

Which in a sense is similar to a probability from the normal distribution.As the name implies the shifted exponential distribution is an exponential distribution that has been horizontally displaced by the shift parameter.In statistical shorthand, which will be used throughout the remainder of the text the ex-Gaussian distribution can be written as

For illustrative purposes, data was simulated first from a normal distribution with $\mu = 10$ and $\sigma = 2$ as well as an exponential distribution with $\lambda = 5$. The sum of values from these data sets is an ex-Gaussian distribution which has skewness

$$\gamma_1 = \frac{2(5)^2}{\left(2^2 + 5^2\right)^{\frac{3}{2}}} \approx 1.60$$

Histogram of these distribution are given in figure 1 realised sample of observations, say $y_1, y_2, \ldots\ldots y_n$, that are possibly positively skewed.
The skewness of a data set can be checked by inspection, using either a histogram or a quantile- quantile (Q-Q) plot, or by estimating the skewness parameter. There is no shortage of software package that can create histograms, so it will not be discussed in this paper. A Q-Q plot is a method of comparing sample data with a known distribution such as the normal distribution. The first step in creating a Q-Q plot is to compute the mean and standard deviation of a sample, namely

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_t \quad \text{and} \quad s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(y_t - \bar{y}\right)^2}$$

The next step is to sort the data in ascending order and create a scatterplot of the values

$$\left(y_1, \phi^{-1}\left(\frac{1}{n+1}\right)\right), \left(y_2, \phi^{-1}\left(\frac{2}{n+1}\right)\right), \ldots \left(y_n, \phi^{-1}\left(\frac{n}{n+1}\right)\right)$$

Where $\phi^{-1}(.)$ is the inverse cumulative distribution function of the standard normal distribution. If the sample data follows a normal distribution, these values will lie in a data follows a normal distribution, this values lie in straight line. A reference line passing through the points

$$\left(\bar{y} + s\phi^{-1}\left(\frac{1}{n+1}\right), \phi^{-1}\left(\frac{1}{n+1}\right)\right) \quad \text{and}$$

$$\left(\bar{y} + s\phi^{-1}\left(\frac{n}{n+1}\right), \phi^{-1}\left(\frac{n}{n+1}\right)\right)$$

is commonly inserted into Q-Q plot . This line represents an exact fit to the normal distribution. For demonstrative purpose , a Q-Q plot was created using 100 observations simulated from an ex-Gaussian distribution with parameters 10, 2 and 5(as above). The simulated data is given in appendix 1.

The scatterplot shown in fig 2 indicates the sample does not follow a normal distribution very well. In fact, the specific curvature of points indicates the data is positively skewed.
The skewness parameter also can be estimated. The methods of moments estimator for skew is

$$\hat{\gamma}_1 = \frac{1}{s^3(n-1)}\sum_{i-1}^{n}\left(y_i - \bar{y}\right)^3$$

For the simulated data set, the estimate of skewness is $\hat{\gamma}_1 \approx 1.46$. since the normal distribution has no skewness, i.e., $\hat{\gamma}_1$ near 0 are indicative of normality and values far from 0 indicate skewness. Brown (1997) gives some discussion regarding skewness; however, there is much debate regarding values of $\hat{\gamma}_1$ that indicate a distribution is skewed(i.e.,how far should $\hat{\gamma}_1$ be from 0 to indicate skewness).

Box-cox Transformation

Box-Cox (1964) introduced a family of transformations as powers of variable *y,* and is often referred to as *power transformations.* The goal is to transform non-normal data to a data set that is reasonably normal. Its simplest form is given by

$$y(v)=\begin{cases}\dfrac{y^{v}-1}{v}, v\neq 0\\ ln(y), v=0\end{cases}$$

Where *v* is chosen to best represent a transformation to the normal distribution. Note that $v=\frac{1}{2}$ and $v=0$ correspond to the square root and logarithmic transformation respectively. A commonly used method is to choose values of *v* from -3 to 3 increments of 0.25(SAS Institute Inc., 2008). Then for each choice of *v* , a regression line is fit through the points Q-Q plot using the method discussed above. The resulting regression lines are then compared using either the coefficient of determination R^2 or the log-likelihood statistic (larger is better in both cases). if the best choice is *v=1,* there is no indication the data should be transformed.

Transforming the data and computing the log-likelihood in each case can be very time consuming . SAS is one of the few statistical packages that will do this for you. SAS uses the PROC TRANSREG procedure to perform a Box-Cox transformation. Sample SAS code using this procedure with stimulated data is given in Appendix 2. Abbreviated results are given in Table1.

The results from table 1 suggest the optimal transformation is $y(-0.5)=-2(y^{-0.5}-1)$. A linear transformation of a normal random variable is also a normal random variable, so, in practice, the constant terms is omitted and the data transformed as $x=y^{-0.5}$.

Log-transformed data is a special case of the Box-Cox transformation (i.e., *v*=0). In this instance, data is transformed as $x_i=ln(y_i)$ for i=1,2,……n. The sample is then analysed using the x_i values. When reporting results from log-transformed data(or data from any other type of transformation), the summary statistics are not easily interpreted on the original scale. Instead, summary statistics computed on transformed scale are back-transformed to the original scale. For instance, summary statistics from log-transformed data are exponentiated (i.e.,$y=e^x$), and summary statistics from square root-transformed data are squared (i.e., $y=x^2$). a statistical analysis using log-transformed data is sometimes referred to as a *geometric mean analysis* because when the sample mean of the transformed data

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{n}\sum_{i=1}^{n}ln(y_i)$$

is exponentiated, the results is the geometric mean, i.e.,

$$\bar{y}_g=e^{\bar{x}}=e^{\frac{1}{n}\sum_{i=1}^{n}ln(y_i)}=\sqrt[n]{x_1.x_2......x_n}$$

The standard deviation from the transformed data, on the other hand, is not interpretable on the original scale because e^s is not an applicable measure of variability on the original scale. In general, data that is presented as $\bar{x}\pm s$ and $\bar{x}\pm s.e.$ is only useful for symmetric data. A much better option is to back-transform confidence limits. For log-transformed data, a (1-α)100% confidence interval computed from the transformed data is

$$\bar{x}\pm t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}$$

Which on the original scale is

$$\left(e^{\bar{x}-\frac{t_{\frac{\alpha}{2},n-1}s}{\sqrt{n}}}, e^{\bar{x}+\frac{t_{\frac{\alpha}{2},n-1}s}{\sqrt{n}}}\right)$$

If a log-transformation is appropriate for the ex-Gaussian data, the summary statistics would be reported as $\bar{y} = 14.4$ and 95% confidence interval (13.41,15.42).

The log-transformation, by design, does not allow for data that can take on negative or zero values since logarithms are undefined on the interval $(-\infty,0]$. However, a data set can be shifted to the right so that all values are positive. This is especially useful for count data which may have 0 values. The instinctive log-transformation in this instance is $x_i = ln(y_i + 1)$ for i=1,2,…..n; however, this shift must be replicated when back-transforming summary statistics where

$$\bar{y}_g = e^{\bar{x}} - 1$$

and

$$\left(e^{\bar{x} - \frac{t_{\frac{\alpha}{2},n-1} s}{\sqrt{n}}} - 1, e^{\bar{x} + \frac{t_{\frac{\alpha}{2},n-1} s}{\sqrt{n}}} - 1 \right)$$

Since the inverse of this transformation is $y_i = e^{x_i} - 1$. The log-transformation for count data has a very nice interpretation in that values that are 0 on the original scale are also 0 on the log scale.

Estimating Ex-Gaussian parameters

As mentioned previously, the ex-Gaussian distribution is defined by the parameters from the normal and exponential distributions, specifically μ,σ and λ. The goal in this type of analysis is to moments or maximum likelihood estimation. The method of moments estimates for the mean, variance and skewness for any data set are $\bar{y}$, s^2 and $\hat{\gamma}_1$ respectively . To find estimated for μ,σ and λ, the method of moments estimates are set equal to mean, variance skewness of the ex-Gaussian distribution,i.e.,

$$\bar{y} = \mu + \lambda$$

$$s^2 = \sigma^2 + \lambda^2$$

$$\hat{\gamma}_1 = \frac{2\lambda^3}{\left(\sigma^2 + \lambda^2\right)^{\frac{3}{2}}}$$

Then these three equations are solved simultaneously for μ,σ and λ in terms of $\bar{y},s^2$ and $\hat{\gamma}_1$. The resulting estimates for the ex-Gaussian parameters are

$$\hat{\mu} = \bar{y} - s\left(\frac{\hat{\gamma}_1}{2}\right)^{\frac{1}{3}}$$

$$\hat{\sigma}^2 = s^2\left[1 - \left(\frac{\hat{\gamma}_1}{2}\right)^{\frac{2}{3}}\right]$$

$$\hat{\lambda} = s\left(\frac{\hat{\gamma}_1}{2}\right)^{\frac{1}{3}}$$

where $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\lambda}$ are estimates of μ,σ and λ respectively. For the simulated data set, the summary statistics are $\bar{y} \approx 15.36, s \approx 6.09$ and $\hat{\gamma}_1 \approx 1.46$ which result in ex-Gaussian parameter estimates $\hat{\mu} \approx 9.87, \hat{\sigma} \approx 2.64$ and $\hat{\lambda} \approx 5.49$. SAS code for these computations is given in Appendix 3.

The maximum likelihood estimates for the ex-Guassisan parameters are more efficient (smaller variance) than method of moments estimates (Heathcote, 1996); however, this method is beyond the scope of this paper since it does not have a mathematical closed form, is computationally intensive and optimisation methods are chosen ad hoc.

In terms of reaction time data, the ex-Gaussian distribution is easily interpretable in that it is the sum of two independent constructs with one having a symmetric distribution (Gaussian) and the other being positively skewed (Exponential). The choice of the normal and exponential distributions is also beneficial in that it allows for more than two constructs to influence reaction time since the sum of independent normal random variables is another normal random variable and the sum of independent exponential random variables is another exponential random variable. In statistical terms , if x_1 and x_2 are independent normal random variables with means μ_1 and μ_2 and variance σ_1^2 and σ_2^2, and independent random variables y_1 and y_2 are exponential random variables with mean λ_1 and λ_2, then if each random variable is pairwise independent the sum of these random variables is an ex-Gaussian distribution, i.e.,

$$x_1 + x_2 + y_1 + y_2 \sim exGaussian\left(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}, \lambda_1 + \lambda_2\right).$$

If two constructs influence reaction time in a symmetric way and another in a positively skewed way, estimating an ex-Gaussian distribution would be appropriate. However, the major drawback to the ex-Gaussian is that it does not lend itself very easily to statistical inference unlike normally distributed data (or data that can be transformed to the normal distribution)., because commonly used significance tests do not exist for the ex-Gaussian distribution. In practical terms, common statistical techniques such as the t- test and ANOVA are not possible with the ex-Gaussian distribution. Distributional assumptions can be made on the ex-Gaussian parameter estimates to create a significance test such as the likelihood ratio test; however, there are many issues estimating these parameters mentioned above which, in the author's opinion, makes any procedure questionable.

Fig (1)

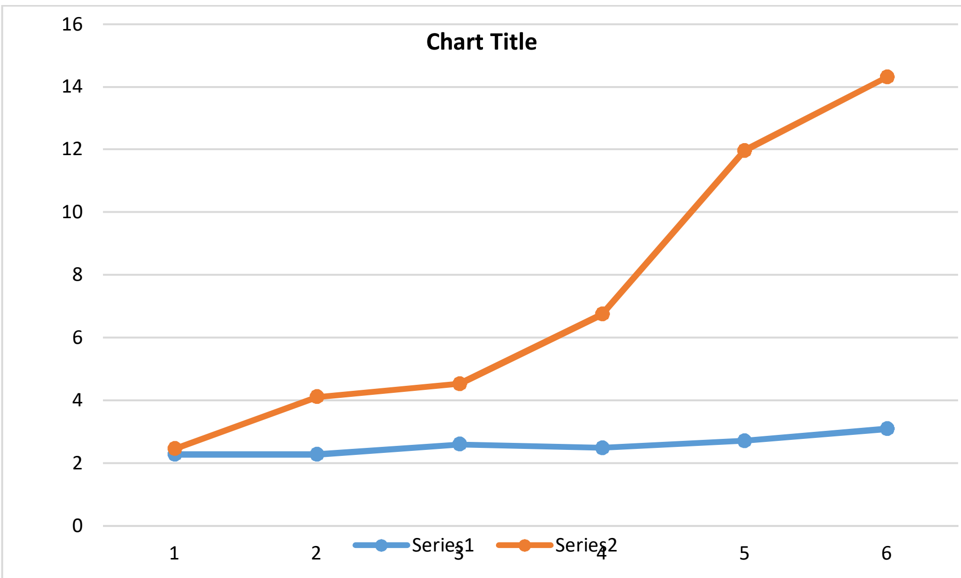

Fig (2)

CONCLUSION

In conclusion, we employed nsPEFs in breast cancer treatment and proved that nsPEFs can inhibit solid tumor growth and tumor blood vessel development. Notably,

nsPEFs exhibited certain selectivity for tumor, as normal skin did not show permanent damage. The data was simulated from an ex-Gaussian distribution, the log-normal appear to be a reasonable approximation to this distribution. This certainly exhaustively Box-Cox transformation is better method for estimating the distribution of a positively skewed data set than estimating the ex-Gaussian distribution. However, the vast world of statistical inference is open to the data analyst if the power transformation reault is reasonable normal data.

REFERENCE

1. Wackerly, D., Mendenhall W., and Scheaffer R. (2007). *Mathematical Statistics with Application* , 7th Edition. Belmont, CA: Duxbury.

2. Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. *Behavior Research Methods, Instruments, and Computers*, 28, 427-445

3. SAS Institute Inc. (2008). SAS/STAT® 9.2 *User's Guide*. Cary, NC: SAS Institute Inc.

4. R. Siegel, D. Naishadham, and A. Jemal. Cancer statitics, 2012, CA. Cancer. J. Clin. 62, no. 1 (2012)10-29

5. L.M. Mir, L.F. Glass, G. Sersa, J. Teissi ~ , C. Domenge, D. Miklavcic, M.J. Jaroszeski et al., Effective treatment of cutaneous and subcutaneous malignant tumors by electrochemotheraphy, Br.J. Cancer. 77, no.12 (1998) 2336

6. T.Goto, T. Nishi, T. Tamura, S. B. Dev, H. Takeshima, M. Kochi, K. Yoshizato et al., Highly efficient electro-gene therapy of solid tumor by using an expression plasmid for the herpes simplex virus thymidine kinase gene, proceedings of the National Academy of sciences 97, no. 1(2000)354-359