

## Implementation of Marker Proteins Using Standardised Effect

Shruti Jain

Jaypee University of Information Technology, Solan, Himachal Pradesh, India.

In this paper we have discussed the null hypothesized mean, standardized effect, true population mean, group sample size, population S.D., type I error rate (Alpha), critical value, and power of different marker proteins (Akt, EGFR, ERK, JNK, MK2, IRS and FKHR) theoretically and validate it with Minitab and Statistica Software. Different plots have been plotted between power and sample size, power and standardized effect, power and type 1 error rate. Later we have plotted a correlation graph for different marker proteins which shows mean, SD, max, min, histogram plot, and scatter plot for each marker protein.

**Keywords:** Marker proteins, Histogram plot, Scatter plot, Power plots.

### Introduction

This paper gives an idea for the different pathways [1-5] which occur due to the of three different inputs i.e. tumor necrosis factor- $\alpha$  (TNF) [6] epidermal growth factor (EGF) [7 -9] and insulin[10, 11]. Out of these seven different marker proteins were predicted those are Akt [12, 13], Epidermal growth factor receptor (EGFR) [14], Extracellular-regulated kinase (ERK) [14 - 17], c-jun N-terminal kinases(JNK) [17- 20], Mitogen-activated protein kinase-activated protein kinase 2 (MK2) [17 - 20], Insulin receptor substrate (IRS), and Forkhead transcription factor (FKHR) [21] which leads to cell death/ cell survival.

**Akt:** Akt leads to cell survival by deactivating different proteins like Raf, Bad, FKHR, caspase 9. PI3 kinase/Akt pathway activates by Phosphatase and tensin homolog (PTEN) [12, 13].

**EGFR :** The epidermal growth factor (EGF) binds with its receptor to form EGF receptor (EGFR) which further binds with Src homology 2 (SH2) [7 -9] leading to the activation of RAS/extracellular signal regulated kinase (ERK) pathway, the phosphatidylinositol 3 kinase(PI3K) pathway, and the activator of transcription (JAK/ STAT) pathway [14].

**ERK :** EGF activates ERK by binding through Grb2/Shc leading SOS which activates RAS leads RAF1 which further phosphorylates to MEK1/ MEK2 which activates ERK1/ ERK2 respectively [14- 17]. It is one of the path-way of MAPK family.

**JNK pathway:** The second pathway of MAP kinase cascade is the stress activated protein kinase/ c-Jun NH<sub>2</sub>-terminal kinase (JNK/SAPK). The JNK is activated when cells are exposed to heat shock [17 -20] and UV radiation.

**p38/ MK2 pathway:** The p38/MK2 is the third and main pathway. It is activated in response to inflammatory cytokines, endotoxins, and osmotic stress [17- 20].

**IRS:** Insulin is a hormone that binds to its receptor (the insulin receptor) on cell membranes and initiates signal transduction leading to cell survival/ apoptosis.

**FKHR:** Binding of Sh2 and ErbB leads to PI3K which is through EGF and insulin receptors [21]. PI3K further leads to phosphorylates phosphatidylinositol (4, 5) biphosphate and RAS. Akt leads to cell death by deactivating FKHR, GSK.

## Implementation Using Standardized Effect

In this paper we have calculated the Null Hypothesized Mean ( $\mu_0$ ), Population S.D. (sigma), Standardized Effect ( $E_s$ ), True Mean Population ( $\mu$ ), Type I Error Rate (Alpha), Power and Critical Value of  $t$  for all marker proteins using 10 different combinations of inputs.

The population mean is expressed as

$$\mu = \frac{\sum X}{N} \quad \text{and} \quad \text{population standard deviation is expressed as,}$$

$$\sigma = \sigma^2 = \sqrt{\frac{\sum (X - \mu)^2}{N}}$$

where  $X$  are the different values,  $\sum X$  is the addition of different values,  $N$  is the population size,  $\sigma^2$  is the population variance. Type 1 error is false positive while Type 2 error is false negative. To have  $p$ -value less than significance level ( $\alpha$  for type 1 and  $\beta$  for type 2), a  $t$ -value must be to the right of  $t_\alpha$ .

There are different types of the Standardized Effect [22]:

### Correlation Family- Effect Based on Variances

**Pearson  $r$  or correlation coefficient:** The value of  $r$  ranges from  $-1$  to  $1$ , where  $-1$  signifies negative relation between two variables,  $1$  signifies positive relation between two variables, and  $0$  signifies linear relation between two variables.

- Coefficient of determination ( $r^2$ ) or squared  $r$ : Coefficient of determination is the square of the Pearson correlation value
- Eta square ( $\eta^2$ ): Eta-square is the ratio of variance for the dependent variable to the total.

$$\eta^2 = \frac{SS_{Treatment}}{SS_{Total}} \quad ..(1)$$

- Omega square ( $\omega^2$ ): The less biased value of the variance is  $\omega^2$  which is expressed as:

$$\omega^2 = \frac{SS_{Treatment} - df_{Treatment} * MS_{error}}{SS_{Total} + MS_{error}} \quad ..(2)$$

for complex analyses  $\eta^2$  is preferable to  $\omega^2$

**Cohen's  $f^2$ :** The  $f^2$  for multiple regression is defined as:

$$f^2 = \frac{R^2}{1 - R^2} \quad ..(3)$$

where  $R^2$  is the squared multiple correlation.  $f^2$  can be given as:

$$f^2 = \frac{\eta^2}{1 - \eta^2} \quad \text{or} \quad f^2 = \frac{\omega^2}{1 - \omega^2} \quad ..(4)$$

**Cohen's  $q$ :** Cohen's is the difference between two Fisher transformed Pearson regression coefficients and given as

$$q = \frac{1}{2} \log \frac{1 + r_1}{1 - r_1} - \frac{1}{2} \log \frac{1 + r_2}{1 - r_2} \quad ..(5)$$

where  $r_1$  and  $r_2$  are the regressions coefficient and the variance of Cohen's is expressed as

$$\text{var}(q) = \frac{1}{N_1 - 3} + \frac{1}{N_2 - 3} \quad ..(6)$$

where  $N_1$  and  $N_2$  are the number of data points in the first and second regression respectively.

A mean effect size  $\theta$  is the ratio of mean of two population ( $\mu_1$  and  $\mu_2$  respectively) to the standard deviation based on either or both the deviations ( $\sigma$ )

$$\theta = \frac{\mu_1 - \mu_2}{\sigma} \quad ..(7)$$

## Difference Family- Effect Size Based on Difference between Means

**Cohen's  $d$**  : Cohen's  $d$  is the ratio of the difference between two means by a pooled standard deviation for the data ( $s$ ), *i.e.*

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s} \quad ..(8)$$

where  $s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$  while the variance for one of the groups is defined as

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_{1,i} - \bar{x}_1)^2 \quad ..(9)$$

and similar for the other group.

**Hedges  $g$** : Cohen's  $d$  is similar to hedge  $g$ , with a difference of scaling factor.

**Glass  $\Delta$** : This is same as that of Cohen's  $d$  except pooled standard deviation ( $s$ ) is replaced by standard deviation ( $s_2$ )

$$\Delta = \frac{\bar{x}_1 - \bar{x}_2}{s_2} \quad ..(10)$$

**Root mean square standardized effect  $\psi$** : The formula for  $\psi$  for one-way ANOVA, is

$$\psi = \sqrt{\left(\frac{1}{k-1}\right) \frac{\sum (\bar{x}_j - \bar{X})^2}{MS_{error}}} \quad ..(11)$$

Distribution of effect sizes based on means.

## Different Power Plots

We have calculated different parameters for different proteins: Null Hypothesized Mean ( $\mu_0$ ) is 0.0000; True Population Mean ( $\mu$ ) is 2.0000; Population S.D. (sigma) is 10.0000; Standardized Effect ( $E_s$ ) is 0.2000; Group Sample Size ( $N$ ) is 25.0000; Type I Error Rate (Alpha) is 0.0500; Critical Value of  $t$  is 2.0639 and Power is 0.1605. These values

© 2009-2017, JGPT. All Rights Reserved.

are used for plotting the power plots. We have plotted different graphs *i.e.* power v/s sample size ( $N$ ) as shown in fig 1 with  $E_s = 0.2$  and  $\alpha = 0.05$ , power v/s standardized effect ( $E_s$ ) shown in fig 2 with  $N = 25$  and  $\alpha = 0.05$ , power v/s type 1 error rate shown in Fig 3 with  $E_s = 0.2$  and  $N = 25$ . Table 1 shows the different parameters used for plotting different plots which were calculated earlier.

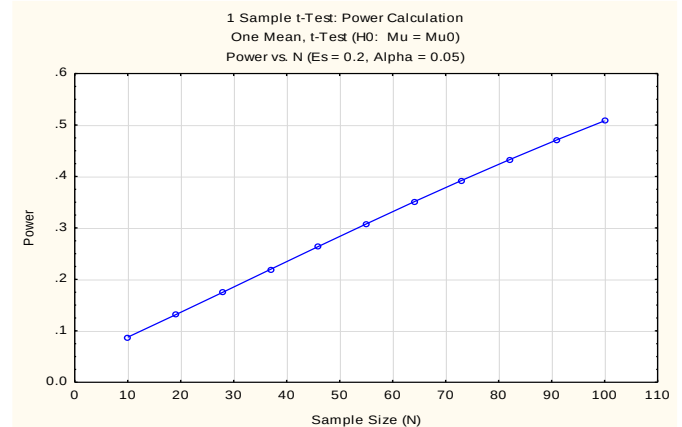


**Fig 1: Plot against power and sample Size ( $N$ )**

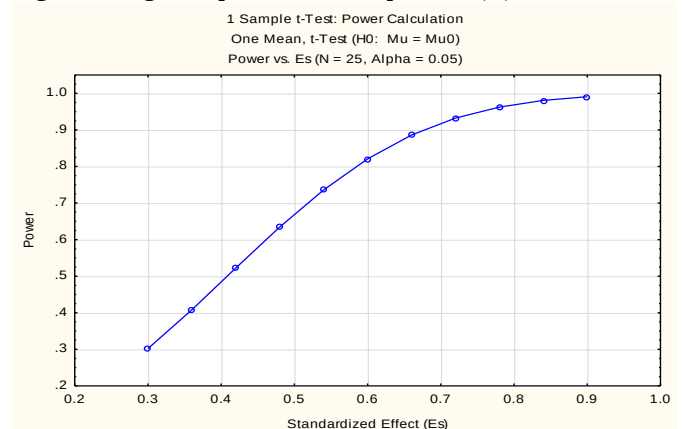


**Fig 2: Plot against power and standardized effect ( $E_s$ )**

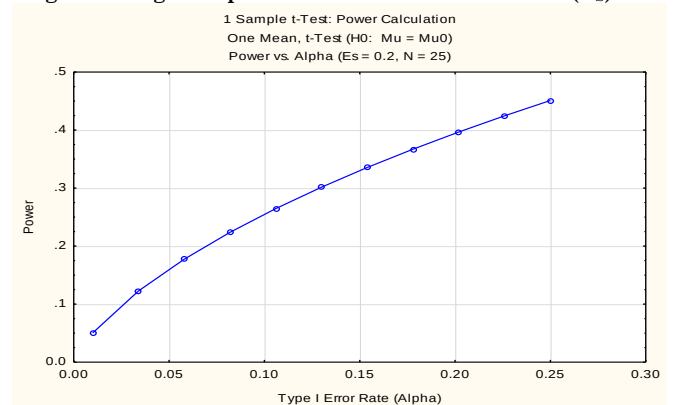


**Fig 3: Plot against power and type I error rate ( $\alpha$ )**

Correlation plots are also shown from Fig 4- Fig 9 ( fig 4 for AkT, fig 5 for EGFR, fig 6 for ERK, fig 7 for FKHR, fig 8 for JNK and fig 9 for IRS) for all marker proteins which shows mean, SD, max, min values of each marker proteins. Scatter plot and histogram are also plotted for all marker proteins.

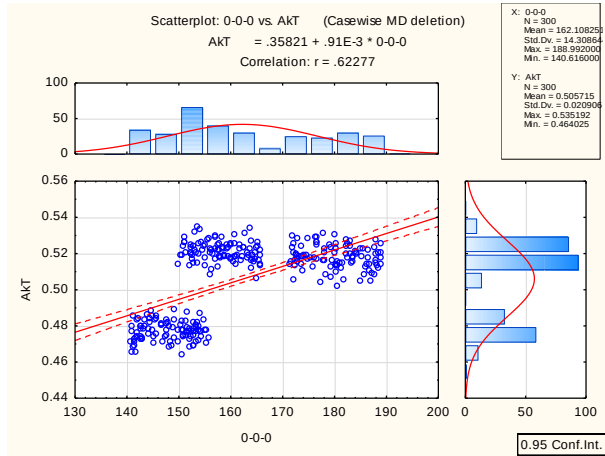


Fig 4: Scatter plot, Histogram , mean, std dev, min and max values for AkT

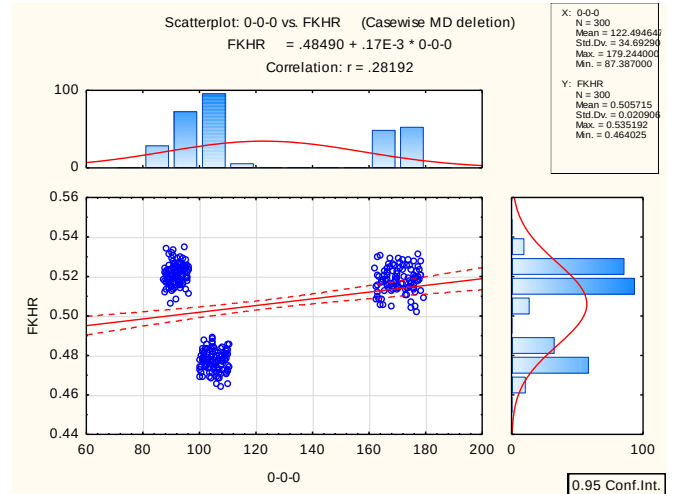


Fig 7: Scatter plot, Histogram , mean, std dev, min and max values for FKHR

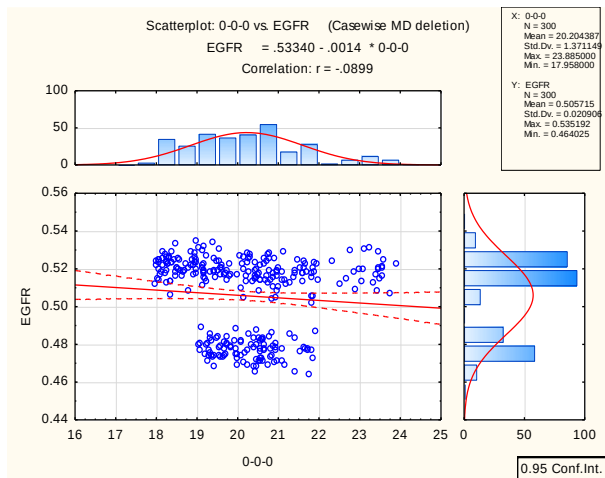


Fig 5: Scatter plot, Histogram , mean, std dev, min and max values for EGFR

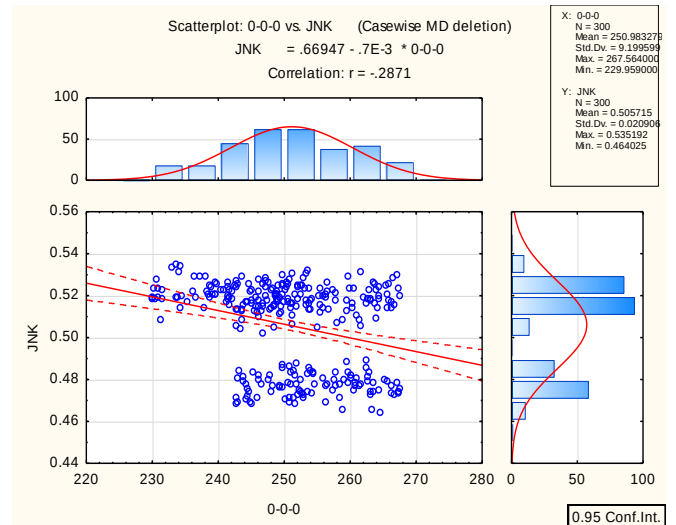


Fig 8: Scatter plot, Histogram , mean, std dev, min and max values for JNK

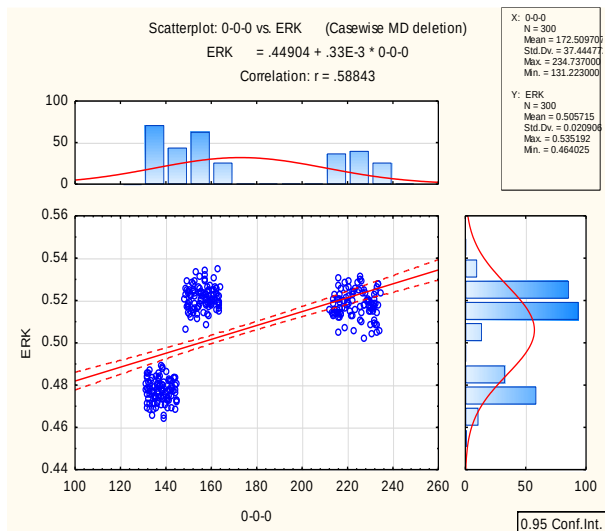


Fig 6: Scatter plot, Histogram , mean, std dev, min and max values for ERK

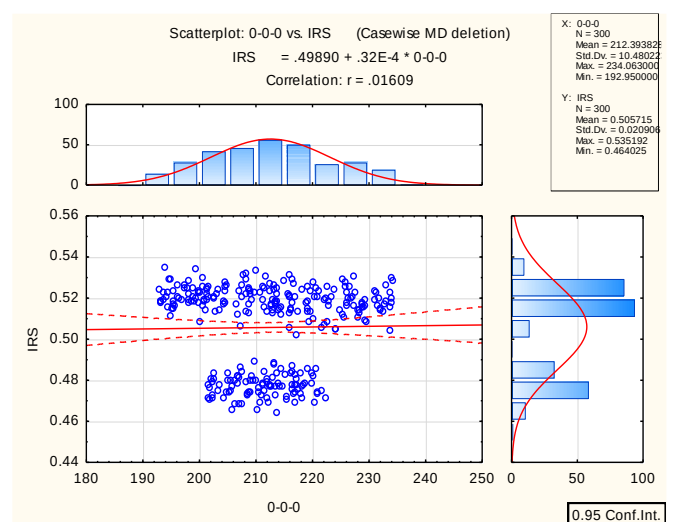


Fig 9: Scatter plot, Histogram , mean, std dev, min and max values for IRS

## Conclusion

In this paper we have calculated the correlation values for all marker proteins. We have plotted scatter plot, histogram for

every marker protein. We have also calculated different parameters like  $\mu_0$ ,  $\mu$ , sigma,  $E_s$ , Alpha, Critical Value of  $t$  and Power for every marker protein.

## References

1. S Jain, (2012) Communication of signals and responses leading to cell survival / cell death using Engineered Regulatory Networks. PhD Thesis, Jaypee University of Information Technology, Solan, Himachal Pradesh, India.
2. S Jain, PK Naik, R Sharma (2009) A Computational Modeling of cell survival/death using VHDL and MATLAB Simulator, Digest Journal of Nanomaterials and Biostructures, 4(4), 863:879.
3. S Jain, PK Naik, (2012) System Modeling of cell survival and cell death : A deterministic model using Fuzzy System, International Journal of Pharma and BioSciences, 3(4):358-373.
4. S Jain, PK Naik, (2012) Communication of signals and responses leading to cell death using Engineered Regulatory Networks, Research Journal of Pharmaceutical, Biological and Chemical Sciences, 3(3):492-508.
5. G Suzanne, AJ Kevin, GA John, AE Pace, AD Lauffenburger, KP Sorger, (2005) A compendium of signals and responses triggered by prodeath and prosurvival cytokines, Manuscript M500158-MCP200, 2005.
6. S Jain, PK Naik, SV Bhooshan (2011) Mathematical modeling deciphering balance between cell survival and cell death using Tumor Necrosis Factor  $\alpha$ , Research Journal of Pharmaceutical, Biological and Chemical Sciences, 2(3), 574-583.
7. TA Libermann, TA Razon, AD Bartal, Y Yarden, J Schlessinger, H Soreq (1984) Expression of epidermal growth factor receptors in human brain tumors, Cancer Res., 44, 753-760.
8. N Normanno, AD Luca, C Bianco, L Strizzi, M Mancino, MR Maiello, A Carotenuto, GD Feo, F Caponigro, DS Salomon, (2006) Epidermal growth factor receptor (EGFR) signaling in cancer, Gene, vol 366, 2:16.
9. S Jain, (2016) Compendium model using frequency/ cumulative distribution function for receptors of survival proteins: Epidermal growth factor and insulin, Network Biology, 6(4), 101:110.
10. M F White (2003), Insulin Signaling in Health and Disease, Science, 302, 171:1711.
11. S Jain, PK Naik, SV Bhooshan, (2011) Mathematical modeling deciphering balance between cell survival and cell death using insulin, Network Biology 1(1):46:58.
12. S Jain, DS Chauhan (2015), Linear & Non Linear Modeling of Protein Kinase B/ AkT, Proceedings of International Conference on Information and Communication Technology for Sustainable Development (ICT4SD - 2015), Ahmedabad, India, , July 3-4, pp 81-88.
13. S Jain, PK Naik, SV Bhooshan, (2012) Compendium Model of AkT for Cell Survival/Death and its Equivalent BioCircuit, International Journal of Soft Computing and Engineering, 2(3):91:97.
14. C. Arteaga, (2003) Targeting HER1/EGFR: a molecular approach to cancer therapy Semin Oncol, 30: 314.
15. S Jain S, PK Naik, SV Bhooshan, (2010) Model of Mitogen Activated Protein Kinases for Cell Survival/Death and its Equivalent Bio-Circuit, Current Research Journal of Biological Sciences, 2(1), 59P:71.
16. S Jain, DS Chauhan, (2015) Mathematical Analysis of Receptors For Survival Proteins, International Journal of Pharma and Bio Sciences, 6(3), 164:176.

17. SJain,(2016)Regression analysis on different mitogenic pathways, Network Biology, 6(2), 40:46.
18. S Jain, (2016)Mathematical Analysis using Frequency and Cumulative Distribution functions for Mitogenic Pathway, Research Journal of Pharmaceutical, Biological and Chemical Sciences, 7(3). 262:272.
19. S Jain (2014) Implementation of Fuzzy System using Operational Transconductance Amplifier for ERK pathway of EGF/ Insulin leading to Cell Survival/ Death, J Pharm Biomed Sci, 4(8):701:707.
20. S Jain, DS Chauhan, (2015)Implementation of fuzzy system using different voltages of OTA for JNK pathway leading to cell survival/ death, Network Biology, 5(2),2015, 62:70.
21. S Jain, (2015)Mathematical Analysis and Probability Density Function of FKHR pathway for Cell Survival /Death, Proceedings of Control System and Power Electronics– CSPE 2015, Bangalore, India. 84:93.
22. [https://en.wikipedia.org/wiki/Effect\\_size](https://en.wikipedia.org/wiki/Effect_size)