Odds and Probabilities Estimation for the Survival of Breast Cancer Patients with Cancer Stages 2 & 3

Jackie D. Urrutia , Ciandreu Noah M. Balmaceda and Edcon B. Baccay
Office of the Academic Head, Polytechnic University of the Philippines- Parañaque Campus
Col. E. L. De Leon St. Sto. Niño Parañaque City, Philippines

Abstract. Breast cancer is one of the leading causes of death in the Philippines. One out of four who are diagnosed with breast cancer die within the first five years, and no less than 40 percent die within 10 years and it has continous rise as time passes by. Therefore, it is very important to know the factors that can help for the survival rate of the patients. The purpose of this study is to identify the best possible treatment or combination of treatments. The researchers considered four independent variables namely: Completed Surgery, Completed Chemotherapy, Completed Hormonotherapy and Completed Radiotherapy. The researchers limit this study for only 160 patients with stage 2 and 135 with stage 3 for a total of 295 patients considering the data gathered from three hospitals from Metro Manila. The names of the hospitals were not declared due to confidentiality of data. In identifying the best treatment or combination of treatments, odds, probabilities and odds ratios of patients, Logistic Regression Analysis was used.

1. Introduction

According to Philippine Breast CancerNetwork (PBCN), the Philippines has the highest incidence rate of breast cancer in Asia and registered the highest increase of 589% among 187 countries over a 30 year period from 1980 to 2010. Breast cancer is the leading killer of women ages 35 to54 worldwide.[1] Also, according to the Philippine Society of Medical Oncology in 2012, an estimated 3 out of 100 Filipino women will contract the disease before age 75 and 1 out of 100 will die before the age of 75.[2]

Breast cancer is a malignant tumor that starts in the cells of the breast. A malignant tumor is a group of cancer cells that can grow into surrounding tissues or spread to distant areas of the body. The disease occurs almost entirely in women, but men can get it, too.[3]

The first symptom of breast cancer that most women notice is a lump or an area of thickened tissue in their breast. Most breast lumps aren't cancerous but it's always best to have them checked by their doctor. An individual should see her doctor if she notices any of the following: a lump or area of thickened tissue in either breast, a change in the size or shape of one or both breast, discharge from either of the nipples, a lump or swelling in either of their armpits, dimpling on the skin of your breast, a rash on or around the nipple and a change in the appearance of the nipple. Breast pain isn't usually a symptom of breast cancer.[4]

Women with breast cancer may have questions about their prognosis and survival. Prognosis and survival depend on many factors. Only a doctor familiar with a person's medical history, type of cancer, stage,

characteristics of the cancer, treatments chosen and response to treatment can put all of this information together with survival statistics to arrive at a prognosis. The researchers considered Completed Surgery, Completed Chemotherapy, Completed Hormonotherapy and Completed Radiotherapy as predictive factors because doctors believe that these factors can affect the survival rate of patients having breast cancer.[5]

Some treatments remove or destroy the disease within the breast and nearby tissues, such as lymph nodes. These include: Surgery, radiation therapy, chemotherapy, hormone therapy and targeted therapy but the researchers only limit the study without the data for targeted therapy due to lack of patients that undergo this treatment.[6]

Surgery aims to remove the breast cancer with a margin (border) of normal tissue to reduce the risk of the cancer coming back in the breast (known as local recurrence) and to try to stop any spread elsewhere in the body. The amount of tissue removed depends on the area of the breast affected and the size of the cancer in your breast. Surgery is usually the first treatment for breast cancer, although sometimes chemotherapy or hormone therapy is offered first. This is to begin treating the whole body or to shrink the cancer so that surgery may be less extensive.[7]

Chemotherapy treatment uses medicine to weaken and destroy cancer cells in the body, including cells at the original cancer site and any cancer cells that may have spread to another part of the body.[8]

Hormone therapy (also called hormonal therapy, hormone treatment, or endocrine therapy) slows or stops the growth of hormone-sensitive tumors by blocking the

body's ability to produce hormones or by interfering with hormone action. Tumors that are hormone-insensitive do not respond to hormone therapy. Hormone therapy for breast cancer is not the same as menopausal hormone therapy or female hormone replacement therapy, in which hormones are given to reduce the symptoms of menopause.[9]

Radiotherapy is a treatment for cancer that uses carefully measured and controlled high energy x-rays. In primary breast cancer it aims to destroy any cancer cells that may be left behind in the breast area after surgery.[10]

Breast cancer is still the leading cause of cancer among women, accounting for 28 percent of the total cases. One out of four who are diagnosed with breast cancer die within the first five years, and no less than 40 percent die within 10 years.[11]Therefore, it is very important to know the factors that can help for the survival rate of patients experiencing breast cancer.

1.1. Objective of the Study

The main objective of the study is to make a model that generates probabilities and odds of surviving a breast cancer using Logistic Regression Analysis. This study also aims to identify which among the following treatments or combination of treatments namely surgery, chemotherapy, hormonal therapy and radiotherapy are significant variables that highly affect the survival rate of breast cancer.

1.2. Statement of the Problem

The study aims to estimate the odds of an individual experiencing breast cancer with factors which are considered to be significant contributors to the survival of the disease using logistic regression. Specifically, it seeks answers to the following questions:

1.2.1. What is the profile of the breast cancer patients for stage 2 & 3?
1.2.2. What is the logistic model for stage 2& 3 cancer that can be obtained to determine the probability of surviving breast cancer?
1.2.3. What are the significant variables that can affect the survival rate of breast cancer patients?
1.2.4. What are the odds and probabilities of an individual of surviving a breast cancer for stage 2 and stage 3 patients?
1.2.5. What are the odds ratios of an individual surviving a breast cancer compare to a breast cancer patient with no treatment applied?
    1.2.5.1. Single Factor
    1.2.5.2. Combination of two factors

1.3. Scope and Limitation

The researchers limited the study as there were only 290 health data of female breast cancer patients with 160 of them are experiencing stage 2 of the disease and 135

patients are with 3rd stage. The data were gathered from three hospitals from Metro Manila. The names of the hospitals were not declared due to confidentiality of the data. The researchers applied logistic regression analysis to identify key determinants of breast cancer from the qualifying health survey data and clinic pathological predictors of the patients.

2. Methodology

2.1. Population and Covariates

The model for odds and probabilities of survival rate of patients were obtained using the data gathered from three hospitals in Metro Manila compiled by the specialists and experts to imply female patients with breast cancer disease. The names of the hospitals were not declared because of the confidentiality of the data. About 95% of the 160 breast cancer patients with stage 2 disease survived the disease and 5% didn't. On patients with 3rd stage, there were 87.4% survived the disease and 12.6% didn't. There were four treatments that contribute to the overall model. These are Completed Surgery, Completed Chemotherapy, Completed Hormonotherapy and Completed Radiotherapy.

2.2. Statistical Tool

The researchers used the statistical software Statistical Package for Social Science (SPSS) in identifying the significant predictors that actually affect the survival of patients as well as computing the odds, probabilities and their odds ratios by using the Logistic Regression. SPSS is one of the most popular statistical packages which can perform highly complex data manipulation and analysis with simple instructions.[12]

2.3. Logistic Regression

Logistic regression is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable in which there are only two possible outcomes.[13]

Logistic Regression does not make many of the key assumptions of linear regression and generalized linear models that are based on ordinary least squares algorithms.

Firstly, it does not need a linear relationship between the dependent and independent variables. Logistic regression can handle all sorts of relationships, because it applies a non-linear log transformation to the predicted odds ratio. Secondly, the independent variables do not need to be multivariate normal – although multivariate normality yields a more stable solution. Also the error terms (the residuals) do not need to be multivariate normally distributed. Thirdly, homoscedasticity is not needed. Logistic regression does not need variances to be heteroscedastic for each level of the independent variables. Lastly, it can handle ordinal and nominal data

as independent variables. The independent variables do not need to be metric (interval or ratio scaled).

However, some assumptions still apply. First, the dependent variable must be binary for binary logistic regression and ordinal for ordinal logistic regression. Second, the outcome variable needs to be coded properly that is, 1 for desired outcome and 0 for the event not occurring. Third, only meaningful variables should be included. Fourthly, independent variables should be independent that is data points should not be from another variable (e.g. before-and-after measurements and matched pairings) and there should be no multicollinearity among the variables. Fifth, logistic regression assumes linearity of independent variables and log odds. Lastly, it requires quite large sample size.[14]

The probability of the outcome is measured by the odds of occurrence of an event. If P is the probability of an event, then (1-P) is the probability of it not occurring.

$$\text{Odds of success} = P / 1\text{-}P$$

The simple logistic model has the form:

$$\text{Logit}(Y) = \text{natural log}(odds) = \ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta X \qquad (1)$$

Taking the antilog of Equation 1 on both sides, one derives an equation to predict the probability of the occurrence of the outcome of interest as follows:

$$\pi = Probability\ (Y = \text{outcome of interest} \mid X = x, \text{a}$$

$$\text{specific value of } X) = \frac{e^{(\alpha+\beta X)}}{1+e^{(\alpha+\beta X)}} \qquad (2)$$

where π is the probability of the outcome of interest or "event", α is the Y intercept, β is the regression coefficient and $e = 2.71828$ is the base of the system of natural logarithms. X can be continuous or categorical, but Y is always categorical.

Extending the logic of the simple logistic regression to multiple predictors, one can construct a complex logistic regression for Y as follows:

$$\text{Logit}(Y) = \ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 +$$
$$\beta_3 X_3 + \cdots + \beta_n X_n \qquad (3)$$

Therefore ,

$$\pi = Probability\ (Y = \text{outcome of interest} \mid X_1 = x_1,$$
$$X_2 = x_2, X_3 = x_3, \ldots X_n = x_n) =$$

$$\frac{e^{(\alpha+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+\cdots+\beta_n X_n)}}{1+e^{(\alpha+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+\cdots+\beta_n X_n)}} \qquad (4)$$

where π is once again the probability of the event, α is the Y intercept, βs are the regression coefficients, and Xs are a set of predictors.[23]

An odds ratio is a measure of association between an exposure and an outcome. The odds ratio represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. Odds ratios are most commonly used in case-control studies; however they can also be used in cross-sectional and cohort study designs as well.[15]

The odds ratio is given by :[16]

$$OR = \frac{[p/(1-p)]_2}{[p/(1-p)]_1} \qquad (5)$$

Where:

$[p/(1-p)]_2$ = odds of success at condition 2
$[p/(1-p)]_1$ = odds of success at condition 1

3. Results and Discussion

3.1. Profile of Patients

Table 1. Profile of Stage 2 & 3 Breast Cancer Patients

The table shows the profile of the breast cancer patients with stage 2 or stage 3 diseases. Majority of the stage 2

Factors	Classification	Stage 2	Stage 3
		Percentage	
Completed Surgery	No	3(1.9%)	27 (20%)
	Yes	157(98.1%)	108(80%)
Completed Chemotherapy	No	32 (20%)	13(9.6%)
	Yes	128 (80%)	122(90.4%)
Completed Hormonotherapy	No	74(46.3%)	46(34.1%)
	Yes	86 (53.8%)	89(65.9%)
Completed Radiotherapy	No	126(78.8%)	43(31.9%)
	Yes	34(21.3%)	92(68.9%)

patients completed a surgery accompanying 98 percent out of 160 while 80% out of 135 patients with stage 3 cancer completed a surgery. Also, majority of stage 2 and stage 3 patients completed a chemotherapy treatment with 80% for stage 2 and 90.4% for stage 3. The hormonotherapy treatment was also completed by 53.8% of stage 2 cancer patients while 65.9% for stage 3 patients. Lastly, only 21.3% of the stage 2 patients completed a radiotherapy treatment and almost 69% for stage patients.

3.2. Logistic Model for Stage 2 Breast Cancer Patients

3.2.1. Significant Parameters

By running the Logistic Regression Analysis, only the completed surgery appears to be significant but the

completed chemotherapy is obviously near the significance level (0.05) which signifies that it is a possible contributor to the model.
The model obtained was:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

$$\ln\left(\frac{p}{1-p}\right) = -2.556 + 4.446 \mathbf{X_1} + 1.863\mathbf{X_2} + 0.092\mathbf{X_3} + 17.508\ \mathbf{X_4}$$

Where:

X_1 = Completed Surgery or Not
X_2 = Completed Chemotherapy or Not
X_3 = Completed Hormonal therapy or Not
X_4 = Completed Radiotherapy or Not

3.2.2. Forward Stepwise Method

Running the Logistic Regression again but this time, removing all the insignificant factors and leaving only all significant variables in the final model (Forward Stepwise Method), the results implied that Surgery and Chemotherapy are the significant variables with regression coefficients of 4.812 and 2.173 respectively.

Odds= exp $(\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4)$
Odds = exp $(-2.866 + 4.812 X_1 + 2.173 X_2)$
The final model obtained was:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2$$
$$\ln\left(\frac{p}{1-p}\right) = \mathbf{-2.866 + 4.812 X_1 + 2.173 X_2}$$

A person's probability of surviving breast cancer with stage 2 is written by:

$$P(Y = 1 | X_1, X_2) = \frac{e^{(-2.866 + 4.812 X_1 + 2.173 X_2)}}{1 + e^{(-2.866 + 4.812 X_1 + 2.173 X_2)}}$$
$$or \ \frac{1}{1 + e^{-(-2.866 + 4.812 X_1 + 2.173 X_2)}}$$

Where:

X_1 = Completed Surgery or Not
X_2 = Completed Chemotherapy or Not

The odds of a stage 2 breast cancer patient who didn't undergo surgery nor chemotherapy is given by $e^{-2.866} = 0.056$ which also means that her probability of surviving breast cancer is Y = (Odds) / (1+Odds) = (0.056) / (1+0.056) = 5.4%.

3.3. Logistic Model for Stage 3 Breast Cancer Patients

3.3.1. Significant Parameters

By running the regression with an alpha of 0.05, the researchers found that among the four initial independent variables, only two are significant: Completed Surgery

and Completed Radiotherapy with regression coefficients of 1.479 and 1.292 respectively.
The model obtained was:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

$$\ln\left(\frac{p}{1-p}\right) = -1.466 + 1.479\mathbf{X_1} + 1.016\mathbf{X_2} + 1.147\mathbf{X_3} + 1.292\mathbf{X_4}$$

Where:

X_1 = Completed Surgery or Not
X_2 = Completed Chemotherapy or Not
X_3 = Completed Hormonal therapy or Not
X_4 = Completed Radiotherapy or Not

3.3.1. Forward Stepwise Method

Removing all insignificant variables and thus leaving just all the significant factors in the final model (Forward Stepwise Method), the researchers concluded that radio therapy is the only treatment that is significant with a regression coefficient of 1.022.
Odds = $\exp(\alpha + \beta_1 X_1)$
Odds = $\exp(1.329 + 1.022 X_1)$

The final model obtained was:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 X_1$$

$$\ln\left(\frac{p}{1-p}\right) = 1.329 + 1.022\mathbf{X_1}$$

A person's probability of surviving breast cancer with stage 3 is written by:

$$P(Y = 1 | X_1) = \frac{e^{(1.329 + 1.022 X_1)}}{1 + e^{(1.329 + 1.022 X_1)}}$$
$$or \ \frac{1}{1 + e^{(1.329 + 1.022 X_1)}}$$

Where:

X_1 = Completed Radiotherapy or Not

The odds of a stage 3 breast cancer patient who didn't undergo radiotherapy is given by $e^{1.329} = 3.777$ which also means that her probability of surviving breast cancer is $Y = \frac{Odds}{1+Odds} = \frac{3.777}{1+3.777} = 79.05\%$.

3.4. Significant Factors

By using the Logistic Regression Analysis, it was concluded that for stage 2 breast cancer patients, the significant predictors are completed surgery and chemotherapy having p-values of 0.001 and 0.015 respectively while on stage 3, radiotherapy is the only significant factor having a p-value of 0.05.

3.5. Odds and Probabilities

The odds of a stage 2 breast cancer patient that competed a surgery only is $e^{-2.866+4.812(1)} = 7.0006$ resulting to $Y = (7.0006) / (1+7.0006) = 87.50\%$ while the odds of a stage 2 breast cancer patient that completed chemotherapy only is $e^{-2.866+2.173(1)} = 0.5001$ resulting to
$Y = (0.5001) / (1+0.5001) = 33.34\%$. The odds of a stage 2 breast cancer patient that undergo both surgery and chemotherapy is $e^{-2.866+4.812(1)+2.173(1)} = 61.4977$ resulting to
$Y = (61.4977) / (1+61.4977) = 98.40\%$.

For stage 3, the odds of a stage 3 breast cancer patient that completed radiotherapy only is $e^{1.329+1.022(1)} = 10.4961$ to $Y = (10.4961) / (1+10.4961) = 91.31\%$.

3.6. Odds Ratio

3.6.1. Odds Ratio of Single and Combination of Treatments

The model predicts the odds ratio for completed surgery to be $OR = [e^{-2.866+4.812(1)}] / (e^{-2.866}) = 122.978$. Therefore, the survival of a stage 2 breast cancer is likely to happen to a patient that has completed surgery by 123 times compared to a person that didn't undergo of any of these treatments. The model also predicts the odds ratio for completed chemotherapy to be $OR = [e^{-2.866+2.178(1)}] / (e^{-2.866}) = 8.7846$. Out of all single and combination factors, the combination of both treatments is better than applying a single treatment only with an odds ratio of more than a thousand indicating that a breast cancer patient that undergo both surgery and chemotherapy will survive the disease by more likely a thousand times compared to a cancer patient that didn't undergo any of these treatments. It is computed as
$OR = [e^{-2.866+4.812(1)+2.178(1)}] / (e^{-2.866}) = 1080.3064$.

3.6.2. Odds Ratio of Single Treatment

The model predicts the odds ratio for completed radiotherapy to be $OR = [e^{1.329+1.022(1)}] / (e^{1.329}) = 2.7790$.
Therefore, the survival of a stage 3 breast cancer is likely to happen to a patient that has completed radiotherapy only by almost thrice compared to a person that didn't undergo any treatment.

4. Conclusions and Recommendations

Researchers conclude that for stage 2, the combination of surgery and chemotherapy are the best treatments that can be applied to a patient for her to get a better chance of surviving the disease and for stage 3, radiotherapy is preferred. The study found out that the survival of the disease for breast cancer patients with stage 2 is more likely to happen to a patient that undergone the combination of surgery and chemotherapy by more than a thousand times compared to a patient that didn't undergone any of these treatments. Also for stage 3, a patient that undergone radiation is three times more likely to survive the disease compared to a patient that didn't undergone any treatment.

The researchers recommend that adding more independent variables such as age, menopausal status, type of cancer, histological grade, lymphatic and vascular invasion, family history and size of primary tumor will enhance the model because some studies proved that these factors really influence the survival rate of patients with breast cancer. The researchers also recommend adding more participating hospitals in order to increase the sample size for a better model.

References

[1] "*Philippine Breast Cancer Network*"(n.d.) Retrieved October 28, 2015 from http://www.pbcn.org

[2] "*PHL has highest breast cancer rate in Asia*"(2013, October 2). Retrieved October 28, 2015 from http://www.gmanetwork.com/news/story/329126/cbb/phl-has-highest-breast-cancer-rate-in-asia

[3] *"What is breast cancer?"* (2015, June 10). Retrieved October 28, 2015 from http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-what-is-breast-cancer

[4] "*Symptoms of Breast Cancer*" (2014). Retrieved November 10, 2015 from http://www.nhs.uk/Conditions/Cancer-of-the-breast-female/Pages/Symptoms.aspx

[5] "*Breast Cancer*" (n.d.) Retrieved October 28, 2015 from http://www.cancer.ca/en/cancer-information/cancer-type/breast/prognosis-and-survival/?region=on

[6] "*Breast Cancer Treatment*" (n.d.) Retrieved December 11, 2015 from http://www.webmd.com/breast-cancer/guide/breast-cancer-treatment

[7] "*Surgery*" (2014) Retrieved December 11, 2015 from https://www.breastcancercare.org.uk/information-support/facing-breast-cancer-going-through-treatment-breast-cancer/surgery

[8] "*Chemotherapy*" (2015) Retrieved December 11, 2015 from http://www.breastcancer.org/treatment/chemotherapy

[9] "*Hormone Therapy for Breast Cancer*" (2012) Retrieved December 11, 2015 from *http://www.cancer.gov/types/breast/breast-hormone-therapy-fact-sheet*

[10] "*Radiotherapy for Primary Breast Cancer*" (2014) Retrieved December 11, 2015 from https://www.breastcancercare.org.uk/information-support/facing-breast-cancer-going-through-treatment-breast-cancer/radiotherapy-primary-breast-cancer

[11] *"The ABCs of breast cancer"* (n.d.). Retrieved October 28, 2015 from http://psmo.org.ph/the-abcs-of-breast-cancer

[12] "*What is SPSS?*" (n.d.) Retrieved November 11, 2015 from https://www.uwindsor.ca/its/sites/uwindsor.ca.its/files/What%20is%20SPSS.pdf

[13] "*Logistic Regression*" (2015, October 22). Retrieved October 28, 2015 from https://www.medcalc.org/manual/logistic_regression.php

[14] *"Assumptions of Logistic Regression"* (n.d.). Retrieved October 28, 2015 from http://www.statisticssolutions.com/assumptions-of-logistic-regression

[15] " *An Introduction to Logistic Regression Analysis and Reporting*" (n.d.) Retrieved November 12,2015 from https://datajobs.com/data-science-repo/Logistic-Regression-%5BPeng-et-al%5D.pdf

[16] "*Explaining Odds Ratios*" (August, 2010). Retrieved November 12, 2015 from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/

[17] Youlden, Danny R., Cramb, Susanna M., Cheng Har Yip & Baade, Peter D. (2014). "*Incidence and mortality of female breast cancer in the Asia-Pacific region*" Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4069805

[18] Hassan MS, Ansari J, Spooner D & Hussain SA. (2010). "*Chemotherapy for Breast Cancer*" Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20878101

[19] P A Ganz, J H Rowland, K Desmond, B E Meyerowitz& G E Wyatt (2005) "*Life after Breast Cancer: Understanding women's health-related quality of life and sexual functioning*" Retrieved from http://jco.ascopubs.org/content/16/2/501.short

[20] DursunDelen, Glenn Walker & Amit Kadam (2005). "*Predicting Breast Cancer Survivability: A comparsion of three data mining methods*" Retrieved from http://www.aiimjournal.com/article/S0933-3657(04)00101-0/abstract

[21] Herman Host M.D. &Eiliv Lund M.D. (2006). "*Age as a prognostic factor in breast cancer*" Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/1097-0142(19860601)57:11%3C2217::AID-CNCR2820571124%3E3.0.CO;2-T/abstract

[22] S.E. Pinder, I.O. Ellis, M. Galea, S.O'Rouke, R.W. Blamey &C.W. Elston (2007). "*Pathological prognostic factors in breast cancer. III. Vascular Invasion: relationship with recurrence and survival in a large study with long-term follow up*" Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2559.1994.tb01269.x/abstract

[23] Graham A. Colditz, MBBS; Walter C. Willett,MD; David J. Hunter, MBBS; Meir J. Stampfer, MD; Joann E. Manson,MD; Charles H. Hennekens,MD; Bernard A. Rosner,PhD& Frank E. Speizer,MD (1993). "*Family History, Age, and Risk of Breast Cancer – Prospective Data from the Nurses' Health Study*" Retrieved from http://jama.jamanetwork.com/article.aspx?articleid=407486

[24] SarojNiraula, Alberto Ocana, Marguerite Ennis &Pamela J. Goodwin (2012). "*Body size and breast cancer prognosis in relation to hormone receptor and menopausal status: a meta-analysis*" Retrieved from http://link.springer.com/article/10.1007/s10549-012-2073-x

[25] Walpole, Ronald E., Myers, Raymond H., Myers, Sharon L., & Ye, Keying (2007) *"Scientists & Engineers guide to probability and statistics"*: Pearson Education, Inc.

[26] "*SPSS Tutorials: Binary Logistic Regression*" Retrieved from https://www.youtube.com/watch?v=Ak_t86zm_sQ

[27] "*Logistic Regression – Interpretation of Coefficients and Forecasting*" Retrieved from https://www.youtube.com/watch?v=cX2sY2La4Ew

[28] "*Logistic Regression – Introduction*" Retrieved from https://www.youtube.com/watch?v=gNhogKJ_q7U

[29] "*Logistic Regression – SPSS (Part 1)*" Retrieved from https://www.youtube.com/watch?v=OvQShzJ7Sns

[30] "*Logistic Regression – SPSS (Part 2)*" Retrieved from https://www.youtube.com/watch?v=zdJhydkcqv4

[31] "*Logistic Regression – SPSS (Part 3)*" Retrieved from https://www.youtube.com/watch?v=hxcDOoupB4Y