1. Introduction:{(2)(7)}
Statistical model is a formal representation for a class of processes that allows a means of analyzing results from experimental studies, such as the Poisson model or the general linear model, it need not propose a process literally interpretable in the context of the individual case.

A model is a statistical description of a physical, chemical or biological state or process. Using a model can help us think about such processes and their mechanisms, so we can design better experiments and comprehend the result. A model forces us to think through (and state explicit) the assumptions behind analyses.

Models can be types as pictorial, conceptual or verbal, physical and mathematical . Consequently, a model is one type of simplified representation of a real system. It describes the systems using mathematical principles the form of an equation or a set of equations. models are developed primarily To solve one or more problems in a real system. As compared to other model types (e.g., pictorial, conceptual and physical), statistical models are often the most useful type.

Regression methods have become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables. It is often the case that the outcome variable is discrete, talking on two or more possible values.

2. Problem Statement:{(1)(9)}
Most statistical models involve unknown parameters that must be estimated from observed data. As new models are created, estimation methods appropriate to those models must be developed. Among the general approaches available for construction of such estimators are the method of moments and the method of maximum likelihood. Method of moments requires the ability to compute population moments of the observable
associated with the postulated models, whereas method of maximum likelihood requires the ability to compute the likelihood of the observed data under the given model. When the model is complex, such computations can prove challenging. In this study the natural history model of breast cancer is of interest in its own right. On the other side, there is a main concept that found in the medical field there is no cure for breast cancer but, there must be more efforts to eliminate this disease or to eradicate. In addition to that, people in many countries still lags behind in the treatment breast cancer in terms of primary health care, medical laboratories, and health centers. That may effect on the high risk of exposure to breast cancer positive without any control or plans to stop the disease from dissemination over all the world among the women. Then, study will attempt to find specified statistical models related to breast cancer.

3. Importance of the Study:
The importance of the study stems from the fact that finding out an appropriate modelling for breast cancer in Sudan will help policy makers conduct a comprehensive strategy to combat this disease thus help improve community health.

4. Study Objectives:
The objectives of this study is:

1. To reveal the concept of breast cancer through statistical models
3. Furthermore, the study aims to test the validity of the model.
4. Determining the factors that affects the breast cancer incidence in Sudan.
5. To apply the complementary log-log regression model so as to discover the relation between variables denotes breast cancer.

5. Study Methodology:
In this study we will apply the statistical models related to the data of the breast cancer. In the study plan, the focusing will be on complementary log-log regression, because it’s suitable to the data collected. Also estimate coefficients and statistical tests can be done to distinguish between the
variables that related to breast cancer incidence and spread through women in the research study.

Data were analyzed by using the Statistical Package for Social Science (SPSS), stata10 and LogXact will be used as required.

6. The Source of Data:
A questionnaire was launched for data collection, patients and controls. In particular this questionnaire, give in Appendix was designed to collect data on risk factors variables. Questions cover all the types of variables. The research depend on female breast cancer patients diagnosed during the period of the research. Areas of control were selected randomly.

Women who are at increased risk for breast cancer can be identified on the basis of their individual risk combination of multiple risk factors or calculation of a women's lifetime probability of breast cancer. Therefore, multivariate risk models have been introduced.

These models allow determination of a woman's composite relative risk for breast cancer as well as her cumulative lifetime risk adjusted individualized breast cancer risk assessment, which is an essential component of the risk-benefit analysis from which decisions regarding the implementation of frequent surveillance, chemoprevention or prophylactic surgery can be made.

Two frequently used models are the Gail and Claus models. Gail et al (1989), described a risk assessment model that focus on non-genetic risk factors with limited information on family history, while Claus et al (1994), assessed the risk level based on family history of breast cancer. A third model, Ford (1994) is based on personal and family history characteristics to identify the presence of any germ-line mutation of the BRCA genes.

There is evidence to support the use of these mathematical models. However, until recently, no single model integrated family history, surrogate measures
of endogenous estrogen exposure and benign breast disease in comprehensive fashion.

A new model, Tyrer-Cuzick is based partly on a dataset acquired from the international breast intervention study (IBIS), (IBIS, 2002). As can be seen, Tyrer-Cuzick model addresses many of the pitfalls of the above models; significantly, the combination of extensive family history, endogenous estrogen exposure, and benign breast disease (atypical hyperplasia).

7.1. Gail Model:

Using multivariate logistic regression, the following risk factors for developing breast cancer were identified in the Breast Cancer Detection Demonstration Project (BCDDP) population: age at menarche, age at first live birth, number of previous breast biopsies, number of first-degree relatives with breast cancer, and current age of the individual. In addition to these characteristics, the demonstration of atypical hyperplasia on biopsy is incorporated into the original Gail model as another multiplication factor. Relative risk estimates were calculated for each of these parameters, and a woman’s composite relative risk is obtained by multiplying the numbers associated with each relative risk factor. Absolute risk defined as the probability of developing breast cancer over a specified time is computed by multiplying the composite relative risk by the baseline proportional hazards estimation derived from the BCDDP population.

Bondy ML et al, (1994) and Costantino JP, Gail MH, Pee D, et al; (1999) stated that the Gail model is routinely used in cancer risk counseling to derive a preliminary breast cancer risk estimate for unaffected women. It is not applicable to women who have already had either in situ or invasive cancers. Although the model has been formally validated in three studies and can accurately predict the rate of breast cancer development in populations, it tends to overestimate risk for young women and underestimate risk for older women. Some of the over prediction in younger women results from the fact
that the model was based on a population of women who were undergoing annual screening mammography.

from the standpoint of genetic risk assessment, the main limitations of the Gail model are that it does not incorporate breast cancer history for more than two first-degree relatives and does not consider age at onset of cancer. Furthermore, because second-degree relatives are not included, paternal family history is ignored. It should also be pointed out that although risk models may be accurate for populations, risk predictions for individuals may be of limited accuracy".

7.2 Claus Model:

A second epidemiologic model used to estimate a woman’s risk of developing breast cancer over time is the Claus model. Claus EB, 1994, using segregation analysis on data obtained from the Cancer and Steroid Hormone Study (CASH), tables were constructed that predict cumulative probabilities for the occurrence of breast cancer at different ages, depending on both the presence of breast cancer in various combinations of first- and second-degree relatives and age at onset of cancer. Although the Claus model is only useful for the subset of women with one or two relatives with breast cancer, it may be more accurate than the Gail model for this cohort, particularly in the setting of premenopausal breast cancer and minor non familial risk factors, and especially when there is a paternal family history of breast cancer.

In general, the Gail and Claus models should be avoided in individuals with a strong family history of cancer and used only with caution when genetic testing has produced negative results.

7.3. Couch Model:

The Couch et al (1997) model is based on data from 169 women who were assessed at a high-risk clinic and tested for mutations in the BRCA1 gene. Risk is based on the average age at diagnosis of breast cancer in a woman’s
family, ethnicity (Ashkenazi Jewish descent or not), the presence of familial breast cancer only or familial breast and ovarian cancer, and whether any individual has had both breast and ovarian cancer.

7.4. Myriad Model:

The Myriad (Frank) model is based on logistic regression analysis of data from 238 women who underwent complete DNA sequencing of the BRCA1 and BRCA2 genes in familial risk clinics across the United States. All of the subjects were diagnosed with breast cancer before age 50 years and had at least one first- or second-degree relative with breast cancer before age 50 or ovarian cancer at any age.

Risk factors include breast cancer status of the proband, two affected sites in the proband (either a second primary breast cancer or ovarian cancer), categorical age of onset in the proband (under 50 years or under 40 years), and breast or ovarian cancer status in a maximum of two relatives. A simple chart lists the probabilities of carrying a mutation for varying combinations of family history.


Risk ratio and risk difference parameters can be related to regression coefficient for fitting binomial data by assuming a functional relationship between disease probabilities and a linear combination of the covariates. This gives multivariate estimation of risk ratio or risk difference parameters while controlling for confounding and considering interaction. The probability $\pi_i$ is functionally related to the linear predictors known as link function. For logistic regression, the link function used is the log of the odds of probabilities; for the risk ratio, the link is $\log(\pi)$ and $\pi=\exp(\beta)$. The log complement link, $\log(1-\pi)=\beta$ described by Weinberg (1986) models the log of the probabilities of no disease, or health. When dealing with chronic conditions it is possible to model binomial data using the complementary log-log link function, $\log(-\log(1-\pi))$, where $\pi$ is the prevalence. In all these
models the regression coefficient $\beta_k$ represents the difference in the probability transformed by the appropriate link function associated with a unit change in the values of the covariates $X_k$ when the other $k-1$ covariates remain constant. For the logistic link $\beta_k$ is the difference between the logarithms of the odds and therefore $\exp(\beta_k)$ is the ratio of the odds associated with a change of one unit in $X_k$. For the log-probability of health link, $\beta_k$ represents the logarithm of the health ratio or ratio of the probability of no disease at values of $X_k$ which differs by one unit. The regression coefficients associated with the covariates in the complementary log-log model estimates the risk ratio and $\exp(\beta_k)$ estimates the rate ratio that approximates the crude prevalence ratio of the disease. In cross-sectional studies, for short follow-up periods, the complementary log-log model is a valid alternative to logistic regression. The fitted values for each link are identical and the parameters (and their estimates) are functionally only when the model is saturated. Since a separate parameter will be available for cell with distinct covariate, the fitted proportions will equal the observed proportions for all links.

Odds and odds ratios were meaningless measures of disease frequency and of exposure effect in cross-sectional studies and in stead the use of prevalences and prevalence ratios was suggested. Others argued that the choice should depend on the health outcome under study, offering examples of valid use of odds and odds ratios in cross-sectional studies. In practice the two measures give almost identical results for low-prevalence diseases; and if interpreted correctly, they will lead to similar conclusions also for non-rare conditions.

Complementary log-log models represent a third alternative to logistic regression and probit analysis for binary response variables. The regression coefficients associated with the covariates in the complementary log-log model estimates the risk ratio and $\exp(\beta_k)$ estimates the rate ratio that approximates the crude prevalence ratio of the disease. In cross-sectional studies, the complementary log-log model is a valid alternative to logistic regression. A transformation of the probability scale that is sometimes useful
in modeling binomial response data is the complementary log-log transformation where \( \log(-\log(1-\pi)) \) is modeled as a linear function of explanatory variables. In the context of a bioassay, the complementary log-log models can be derived by supposing that the tolerances of the individuals have an extreme value distribution known as the Gumbel distribution with probability density function:

\[
f(u) = \frac{1}{k}e^{(u-\alpha)/k} \exp(-e^{(u-\alpha)/k}), \quad -\infty < u < \infty \tag{1}
\]

where \(-\infty < \alpha < \infty\) and \(k > 0\) are known parameters.

Unlike the normal and logistic distributions, on which the probit and logistic transformations are based, this is an asymmetric tolerance distribution with a mode at \(\alpha\) and a mean that is greater than \(\alpha\). The probability of a response when exposed to dose \(d_i\) is

\[
P_i = \int_{-\infty}^{d_i} f(u)du = 1 - \exp[-e^{(d_i-u)}] \tag{2}
\]

And so:

\[
\log(-\log(1-p)) = \beta_0 + \beta_1 d \tag{3}
\]

Where \(\beta_0 = -\alpha/k\) and \(\beta_1 = 1/k\) are unknown parameters.

The complementary log-log function again transforms the probability in the range (0, 1) to \((-\infty, \infty)\) but this function is not symmetric about \(p=0.5\). A Gumbel tolerance distribution is seen in problems concerning the breaking strength of materials. The weakest link theory, according to which a component breaks when the applied force is greater than the tolerance of the weakest link, suggests that a distribution of minima, such as the Gumbel distribution, should be used to model the breaking strength. This leads to a complementary log-log model for the probability of the component failing. Logit and probit models are not appropriate when \(\pi\) increase from 0 fairly slowly but approaches 1 quite
suddenly. Complementary log-log models are frequently used when the probability of an event is very small or very large.

The response curve is defined as:

\[ \pi(x) = 1 - \exp(-\exp(\alpha + \beta x)) \] and \[ \log[-\log\{1-\pi(x)\}] = \alpha + \beta x \] \hspace{1cm} (4)

Is called the complementary log-log link, unlike logit and probit the complementary log-log function is asymmetrical.

A graph of the complementary log-log function is given in Fig (1):

![Complementary Log-log Function](image)

*Fig (1) Complementary log-log function*

Complementary log-log function is asymmetric and rather different from the others (although it is barely distinguishable from the logistic at small values of \( \pi_i \)). According the Charners and Cox (1967), it is difficult to distinguish between the probit and logistic models unless the sample size is extremely large. Ashford and Sowden (1970) generalized the probit model for
multivariate binary responses. In principle any of these functions can be used as the basis for a link function in a generalized linear model.

**9. DATA ANALYSIS AND APPLICATION**

**9.1. Introduction:**

This study was planned to study the epidemiology of breast cancer among Sudanese women. After a thorough discussion with oncologists, an interview schedule was prepared. The structured questionnaire contained the questions regarding all the variables that could be related to breast cancer risk and should be studied for the population of Sudan. For the design of the study, a case-control study was planned. The case-control study is a primary tool for the study of factors related to disease incidence. The data for patients were collected from the Radiation and Isotopes Center (RICK) in Khartoum state, target population is all population above 20 years in Khartoum state during the period of survey. The data were scrutinized to confirm whether the patients attending the hospital were diagnosed as breast cancer cases. The same structured questionnaire was used to interview the cases and controls. The questionnaire was filled in personally from the respondents.

**9.2. The Risk Factors for Female Breast Cancer:**

The areas in Sudan where cancer is prevalent are, Northern States, Khartoum, Gezira, Kordufan, Darfor and the other cases from other stat as Southern and Eastern states. Most of these patients come from Northern and Eastern States. The most common cancer for female in Sudan is breast cancer (34.5%) and esophageal cancer in male (11%). Because these numbers are reaching epidemic proportions, it is vastly important topic to consider (RICK, 2003). From (WHO) in 2005 cancer killed approximately 22000 people in Sudan 17000 of these people were under the age 70. In 2002 breast cancer is the most common cancers found in women in Sudan and in 2005 breast cancer is the leading cause of cancer deaths in women in Sudan.
Any women may develop breast cancer. However, the following risk factors may increase the likelihood of developing the disease. Both genetic and environmental factors are believed to play a role in a woman's risk of developing breast cancer. If either a woman's mother or sister has breast cancer, the woman's risk increases about two to three times. Having both a mother and a sister with breast cancer increases a woman's risk up to six-fold. If that relative had bilateral breast cancer or was diagnosed at an early age, the risk may be further increased. In small groups of families, the patterns of breast cancer incidence seem to be consistent with known patterns of genetic inheritance.

The common risk factors of breast cancer is, gender, age, menarche and menopause, pregnancy, induced abortion, oral contraceptive use, breast feeding, weight control, radiation exposure, Reproductive and ministerial factors, Hormone replacement therapy, History of biopsy and benign breast disease, Family History, Inheritance, and Breast Cancer Risk, Breast Density, Job Strain, Genetics Mutations and anxiety.

9.3 The Patients Sample Size: \{(14)\}

Using the record of the infected women of breast cancer in the selected period we can select a relative sample by using the statistical formula of sample size. Since the sample designed required is to select simple random sampling, the estimation of the incidence rate can be with absolute precision, the (Cochran's formula) for the sample size is given by:

\[
n_0 = \frac{r^2 pq}{d^2} \quad \text{.................................(5)}
\]

In this research we used \( p = 18\% \) as the incidence of breast cancer estimated from hospital records (RICK, 2007). Then when we use the the formula (5) we find that, \( n_0 = 227 \) patients.
Cochran’s (1977) correction formula should be used to calculate the final sample size. These calculations are as follows:

\[
n = \frac{n_0}{1 + \frac{n_0}{\text{population}}} = 165 \quad \text{…………………………………… (6)}
\]

Where:

\(n_0\) = required return sample size = 227

Population of breast cancer treated at (RICK) = 615 patients.

Then the sample size for this study is 165 patients.

9.4 The Results of the Complementary Log-Log Regression Model

Complementary Log-Log Regression analysis of the data was carried out to obtain model describing the relationship between response variable and independent variables. The regression coefficients associated with the covariates in the complementary log-log model estimate the rate ratio. In cross-sectional studies, for short follow-up periods. The two measures namely, odds ratios and prevalence ratios, give almost identical results for low prevalence diseases and if interpreted correctly, they will lead to similar conclusions also for none-rare conditions.
Table (1)

The Contribution Of The Variables In The Complementary log-log Models

<table>
<thead>
<tr>
<th>Effect</th>
<th>Deviance</th>
<th>Change in Deviance</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Feeding</td>
<td>436.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Residence</td>
<td>435.1</td>
<td>1.20</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Race</td>
<td>360.0</td>
<td>75.1</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Occupation</td>
<td>346.5</td>
<td>13.5</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Education level</td>
<td>297.5</td>
<td>49.0</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Marital status</td>
<td>291.4</td>
<td>6.10</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Family history</td>
<td>289.5</td>
<td>1.90</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Menstrual begin</td>
<td>289.4</td>
<td>0.10</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Parity</td>
<td>288.2</td>
<td>1.20</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Menopause status</td>
<td>258.3</td>
<td>29.9</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>Contraceptive use</td>
<td>258.3</td>
<td>0.00</td>
<td>.0000 &lt; .01</td>
</tr>
<tr>
<td>First pregnancy</td>
<td>247.9</td>
<td>10.4</td>
<td>.2845 &gt; .05</td>
</tr>
<tr>
<td>Breast density</td>
<td>239.6</td>
<td>8.30</td>
<td>.7608 &gt; .05</td>
</tr>
<tr>
<td>HRT</td>
<td>208.8</td>
<td>30.8</td>
<td>.9941 &gt; .05</td>
</tr>
<tr>
<td>Benign breast</td>
<td>204.5</td>
<td>4.30</td>
<td>.9977 &gt; .05</td>
</tr>
<tr>
<td>BMI</td>
<td>193.9</td>
<td>10.6</td>
<td>.9981 &gt; .05</td>
</tr>
<tr>
<td>Anxiety scale</td>
<td>193</td>
<td>0.90</td>
<td>.9997 &gt; .05</td>
</tr>
</tbody>
</table>

Source: SPSS, LogXact output

Multivariate models were applied to this case-control study by using complementary log-log functions. One additional covariate was included in every next model. Using the customary approach in regression analysis, change in deviance was used to measure the contribution of the variables in the model by adding a new variable at each successive stage. The results for each successive stage were presented in table (3.24). Multivariate model was obtained using complementary log-log link and compared with the model obtained by using logit link for the same case control data set. factors considered in the model were, breast feeding, race, occupation, education.
level, parity, menopause status, contraceptive use, first pregnancy, breast density, hormone replacement therapy (HRT) and body mass index (BMI).

Table (2)
The Estimated Coefficients of Complementary log-log Regression Model For Case-Control Study

<table>
<thead>
<tr>
<th>Effect</th>
<th>B</th>
<th>S.E</th>
<th>Sig</th>
<th>95% CI of OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Feeding</td>
<td>.6376</td>
<td>.4575</td>
<td>.1634</td>
<td>-.2591, 1.534</td>
</tr>
<tr>
<td>Residence</td>
<td>-.3629</td>
<td>.2523</td>
<td>.1503</td>
<td>-.8573, .1316</td>
</tr>
<tr>
<td>Race</td>
<td>1.790</td>
<td>.4269</td>
<td>.0000</td>
<td>.9532, 2.627</td>
</tr>
<tr>
<td>Occupation</td>
<td>.7058</td>
<td>.3423</td>
<td>.0392</td>
<td>.0349, 1.377</td>
</tr>
<tr>
<td>Education level</td>
<td>-.9417</td>
<td>.1843</td>
<td>.0000</td>
<td>-1.303, -.5804</td>
</tr>
<tr>
<td>Marital status</td>
<td>-.5194</td>
<td>.2313</td>
<td>.0247</td>
<td>-.9728, -.0660</td>
</tr>
<tr>
<td>Family history</td>
<td>.7430</td>
<td>.3455</td>
<td>.0315</td>
<td>.0658, 1.420</td>
</tr>
<tr>
<td>Menstrual begin</td>
<td>.0635</td>
<td>.2951</td>
<td>.8297</td>
<td>-.5149, .6418</td>
</tr>
<tr>
<td>Parity</td>
<td>-.0737</td>
<td>.1915</td>
<td>.7003</td>
<td>-.4491, .3017</td>
</tr>
<tr>
<td>Menopause status</td>
<td>.9563</td>
<td>.2814</td>
<td>.0007</td>
<td>.4048, 1.508</td>
</tr>
<tr>
<td>Contraceptive use</td>
<td>.4842</td>
<td>.2933</td>
<td>.0988</td>
<td>-.0907, 1.059</td>
</tr>
<tr>
<td>First pregnancy</td>
<td>.4093</td>
<td>.1330</td>
<td>.0021</td>
<td>.1487, .6699</td>
</tr>
<tr>
<td>Breast density</td>
<td>-.5083</td>
<td>.2431</td>
<td>.0365</td>
<td>-.9848, -.0317</td>
</tr>
<tr>
<td>HRT</td>
<td>1.635</td>
<td>.3208</td>
<td>.0000</td>
<td>1.007, 2.264</td>
</tr>
<tr>
<td>Benign breast</td>
<td>-.8404</td>
<td>.3829</td>
<td>.0282</td>
<td>-1.591, -.0899</td>
</tr>
<tr>
<td>BMI</td>
<td>-.5033</td>
<td>.1608</td>
<td>.0017</td>
<td>-.8184, -.1881</td>
</tr>
<tr>
<td>Anxiety scale</td>
<td>-.2447</td>
<td>.2623</td>
<td>.3510</td>
<td>-.7589, .2695</td>
</tr>
</tbody>
</table>

Source: SPSS, LogXact output

According to the results of the complementary log-log model, a significant increase in risk of breast cancer was observed for race, occupation, education level and marital status. Family history of breast cancer was a significant risk factor for this research. Sudanese females menopausal status and first full term of pregnancy were at significantly increased risk of breast cancer.
Breast density, HRT and obesity (BMI) also were at significantly increased risk of breast cancer. As shown in table (3.19) and (95% confidence interval) for the variables was obtained from a logistic regression model and compared with the same result shown in table (3.25) obtained from a complementary log-log link. Similar pattern of comparison was observed for each covariate between the crude prevalence ratio and the prevalence ratio estimated by fitted values of the model. Both the ratios were almost similar.

9.5. Statistical and Biological Results:
Model in this study were based on pure statistical results. However the biological aspects of breast cancer are also equally important but not considered earlier. It would not be appropriate to suggest a statistical model in isolation from biology of the disease. The appropriate model would be which provides best result both statistically and reasonably supported biologically. The statistical results that were not supported biologically were revised accordingly in order to obtain a model that can meet both statistical and biological standards.

Complementary log-log transformation is a valid alternative to logistic functions. However confidence intervals of the rate ratio should not be the choice due to asymmetrical nature of the complementary log-log function.

According to the results of the complementary log – log model, a significant increase in risk of breast cancer was observed for family history of breast cancer, menopause status, first full Term pregnancy, breast density, HRT, benging breast and body mass index. Marital status was a significant risk factor for this model. Also race, occupation and education level were statistically significant in this model.

10. CONCLUSIONS AND RECOMMENDATIONS
10.1. Conclusions:
The technique of complementary log-log regression is most commonly used for analysis of epidemiological studies. The complementary log-log function
has been applied to this case–control study and the results obtained from various models are compared.

From the results, residence and anxiety scale statistically is not significance which it means these variables can not play an important role to develop breast cancer among Sudanese women.

Finally in this study model constructed for Sudanese females breast cancer and the model included the independent variables age at birth of first live child, family history of affected mother or sisters, number of previous benign breast biopsy examinations, weight (BMI) and breast density (DENSITY) as dependent variable.

10.2. Recommendations:
The study has the following recommendations:

- Complementary log-log transformation is a valid alternative to logistic functions irrespective of the size of the study.
- In principle, the case – control data used in this study, together with cancer registry data, can be used to construct such models of absolute risk.
- models with such factors might be useful for identifying changes in exposures that might reduce the absolute risk of breast cancer and might be used in counseling.
- The interrelation between breast cancer incidence and risk factors of breast cancer need to be evaluated by the other studies.
- The awareness about the dangerous of the disease still very weak so community, government, civil society, and non-governmental organization they should play an important role to raise the awareness.
- Early detection should be integrated into primary health care
- Early diagnosis and screening should be carried out, specifically targeted towards breast cancer.
- Anxiety test should be done for all women before they diagnosed as a breast cancer women for more reality.
References


[2]. Christopher. B. S (2006). Introduction to mathematical modeling of crop growth, how the equation are derived and assembled into a computer model, PhD, Dissertation, Faculty of Agriculture, University of Putra Malaysia.


[8]. Frank,G(1980).Risk factors in breast cancer:are they important, are they the same in pre- and post menopausal breast cancer patients? Oncology 37:41-5.


