CHAPTER ONE

INTRODUCTION

:Preface .1.1

Statistics is a mathematical science, which has wide applications in various academic disciplines. Within the

natural sciences and the social sciences and humanities, and even government policy.

Among these applications applied to a wide range of topics in the science of medicine and the life sciences.

Survival analysis is widely applied in many fields such as biology, medicine, public health, and epidemiology, where the dependent variable is the time until the event.

Cancer is a public health problem globally. As it represents the most common causes of death (WHO 2008). One of the cancers is breast cancer that has been reported to be the most prevalent cancer among women and accounts for 21.4% of all malignancies. In developed countries breast cancer is detected in the early stages and in more phases of pre-cancer. While in developing countries are still large numbers of patients are diagnosed at an advanced stage in the breast or spread in other parts of the body, and where early detection of the tumor reduces the risk of death. A typical analysis of survival data involves the modeling of time-to-event data, such as the time until death. The time to the event of interest is called either survival time or failure time.

1.2. **Research Problem:**

Survival analysis is another method used in the analysis of data from intervention trials, cohort studies and data routinely collected by cancer registries. It is particularly useful when the probability of occurrence of the event under study changes with time since entry into the study.

Analysis of survival is one of the most important applications in Biostatistics. Scientific studies have uncovered a number of risk factors for breast cancer. Some of these risk factors can be modified by individuals to lower their risk, and others cannot. These factors reduce the hazard ratio. thus has an effect in the survival of breast cancer patient for alive long time. When Cox model to estimate the hazard ratio we know the hazard ratio for each factor. And the application of these models help to identify the characteristics that lead to an increased probability of survival.

1.3. **Research Importance:**

This research is a few of the research applications in the field of bio-statistics, and also compare the survival time among patients who given chemotherapy and those who did not given chemotherapy, as well as to reach a Cox model, which includes the most influential factors in the survival for breast cancer patients.

1.4. Research Objectives:

The Kaplan Meier procedure will use to find the percentiles of survival at any time of interest and to compare the survival time of two studied groups. Cox regression will use to examine the effects of continuous covariates. Also other objectives is to:

1- measure the percentiles survival time after the disease

2- Comparison between the hazard function, for the different disease stages.

3- Comparison between the hazard function, for different types of treatment.

1.5. Research Hypothesis:

1- There is no significant differences on hazard ratio among patients who given the chemotherapy and who did not given chemotherapy.

2- There is no significant difference to the hazard ratio in terms of the stage of the disease.

3- There is no significant differences between patients who given radiotherapy and those who did not given radiotherapy in terms of hazard ratio.

4- There is no significant differences between patients who given hormonal therapy and those who did not given hormonal therapy in terms of hazard ratio.

5- There is no significant differences on hazard ratio among patients who have undergone surgery, and those who did not undergo surgery.

1.6. Data Sources:

Data were collected from the National Center for radiotherapy and nuclear medicine in Khartoum, at period from 2009 to 2011.

1.7. Research Methodology:

Descriptive method to describe research data and analytical method inferential to study survival analysis. .Will use statistical software SPSS, STATA and EXCEL

:Previous studies .1.8

1. In (2009), the researchers Jamal Eivazi Ziaei , Zohreh Sanaat, Iraj Asvadi, Saeed Dastgiri, Ali Pourzand and Jalil Vaez Publish a scientific paper in Asian Pacific Journal of Cancer Prevention entitled "Survival Analysis of Breast Cancer Patients in Northwest Iran", Objective: The objective was to examine survival rates in Tabriz (Northwest of Iran) and comparing with those of data reported from other cities and countries. Results: Survival analysis demonstrated a lower survival rate compared to western countries. Conclusions: Survival rates for our patients are similar/better than other cities in Iran, but lower than certain European countries and the US.

.2 In (2009), The researchers Anjali D. Deshpande, Donna B. Jeffe, Jennifer Gnerlich, Ayesha Z. Iqbal, Abhishek Thummalakunta, and Julie A. Margenthaler Publish a scientific paper in Journal of Surgical Research entitled " Racial Disparities in Breast Cancer Survival: An Analysis by Age and Stage". Results: In the 1988–2003 Surveillance, Epidemiology, and End Results data, 20,424 Black and 204,506 White women were diagnosed with first primary breast cancer. In unadjusted models, Black women were more likely than White women to die from breast cancer (HR: 1.90; 95% CI: 1.83–1.96) and from all causes (HR: 1.52; 95% CI: 1.48–1.55) during follow-up. In models stratified by age and stage, Black women were at increased risk of breast-cancer-specific mortality within each stage group among women <65 years. Conclusions: Racial disparities in breast-cancer specific mortality were predominantly observed within each stage at diagnosis among women <65 years old. This greater mortality risk for Black women was largely not observed among women .>65 years of age

.3 In (2010), The researchers K. Arkoob, M. Al-Nsour, O. Al-Nemry and B. Al-Hajawi Publish a scientific paper in Eastern Mediterranean Health Journal entitled " Epidemiology of breast cancer in women in Jordan

patient characteristics and survival analysis", the study included 838 women . The overall Kaplan–Meier 5-year survival rate was 59.3%. Stage, laterality and grade had a significant effect on survival rate. the strength of the study was the ability to investigate the impact of demographic, histological and therapeutic factors on survival in breast cancer. Furthermore, the size of the study was sufficiently large to perform survival analysis across different .subgroups of breast cancer cases

.4 In (2012), The researchers Nazera Khalil Dakhil, Yahya Mahdi Al–Decemberali and Muna Abbas Mseer Al–A'bidy Publish a scientific paper in Journal of Kufa for Mathematics and Computer entitled "Analysis of Breast Cancer Data using Kaplan–Meier Survival Analysis", Objective The aim of this research was mainly concerned with a study and analysis an estimation of the survivorship time of real data of breast cancer patients in Iraq. Results: the survival experience of benign tumor group is more favorable than the survival experience of malignant tumor group and other tumors groups. Conclusions: With the Kaplan–Meier survival analysis procedure, you have examined the distribution of time to effect for two or more different groups. The comparison tests show that there is a statistically significant difference in survival times (P<5%) .between malignant and benign tumors group only

.1.9 **:Research Organization**

:The research includes four chapters

The **First** chapter contains: Preface, problem, Importance, objectives, assumptions, methodology, research limits , and previous studies. Chapter **Two** contains: Theoretical framework for research**.** Chapter **Three** contain the .practical side of the research

The **Fourth** chapter includes the results reached from the analysis of the research, and the conclusions and .recommendations, and then references and appendices

CHAPTER TWO

BREAST CANCER

:Preface .2-1

Breast cancer is the most important types of cancer for females, which is the most common species for them. Studies indicate that one woman out of every eight women susceptible to breast cancer in the period of her life.

:Cancer .2-2

Cancer begins when cells in a part of the body start to grow out of control. There are many kinds of cancer, but they all start because of this out-of-control growth of abnormal cells. Cancer cell growth is different from normal cell growth. Instead of dying, cancer cells keep on growing and form new cancer cells. These cancer cells can grow into (invade) other tissues, something that normal cells cannot do. Being able to grow out of control and invade other tissues are what makes a cell a cancer cell.

:Breast cancer .2-3

Breast cancer is a malignant (cancer) tumor that starts in the cells of the breast. It is found mostly in women, but men can get breast cancer, too. Although the etiology of breast cancer is unknown, numerous risk factors may influence the development of this disease including genetic, hormonal, environmental, sociological and physiological factors. Over the past few decades, while the risk of developing breast cancer has increased in both industrialized and developing countries by 1%–2% annually, the death rate from breast cancer has fallen slightly.

:Causes of breast cancer .2-4

Though the exact causes of breast cancer are largely unknown, research has found some probable causes of breast cancer.

Family history has long been known to be a risk factor for breast cancer. Both maternal and paternal relatives are important. The risk is highest if the affected relative developed breast cancer at a young age, had cancer in both breasts, or if she is a close relative. First-degree relatives, (mother, sister,

daughter) are most important in estimating risk. Several second-degree relatives (grandmother, aunt) with breast cancer may also increase risk. Breast cancer in a male increases the risk for all his close female relatives. Having relatives with both breast and ovarian cancer also increases a .woman's risk of developing breast cancer

:Some other probable causes and risk factors are

1- .Advancing age
2- .Excessive exposure to radioactive rays
3- .Hereditary genes or family history
4- .Late childbearing
5- .The use of hormone replacement therapy
6- Early onset of a menstrual cycle and an early .menopause
7- .Men or women working in chemical factories

:Stages of breast cancer .2-5

There are several ways to divide the tumors so the doctor can determine the stage of the disease and then give the appropriate treatment according to the stage. And these divisions are generally dependent on three factors: Tumor size, Lymph Nodes and Metastasis and denoted by TNM, (D. .(Mahmoud Shaheen and others

When you get all this information we can divide the breast :tumors into five stages

:Stage 0

At this stage, the cancer is localized, which is very early cancer in the breast does not invade neighboring cells, and can be eradicated and keep the breast or mastectomy as a whole and . denoted by

:Stage I

Is an early stage of breast cancer and may affect the adjacent tissue, which means that the cancer in first stage did not . exceed breast and denoted by

:Stage II

Is also an early stage of breast cancer may affect the tissue adjacent the cancer has spread to the lymph nodes under the armpit may be on two levels. Stage IIA or Stage IIB and .denoted by or or or

:Stage III

Is the stage of the cancer localized Advanced, and have spread to more lymph nodes under the armpit, and perhaps in other

tissues adjacent to the breast. They may be 3 degrees Stage . IIIA or Stage IIIB or Stage IIIC and denoted by or or or

:Stage IV

It stage of Metastasis, cancer move from breast to the rest of . the body as the bone, lung, liver and brain and denoted by

:Breast cancer treatment .2-6

:the treatment of breast cancer Depends on

.Stage of disease .1

. The quality of cancer cells .2

.The patient's desire .3

:Methods of Treatment

: Radiotherapy -1

It is use of high-energy rays to kill cancer cells and prevent them from growing. Radiation be either external radiation .or implant radiation

:Given radiation therapy in two cases

the first case to be complementary to surgical - 1 .treatment

The second case is given when tumor is in areas not - 2 .recommended to eradication the tumor surgically

:Chemotherapy -2

Is the use of medications and drugs to kill cancer cells and in the majority of cases treated breast cancer a variety of medications, given medicines, either by mouth or by injection into a vein or muscle, and in all ways is a chemotherapy treatment comprehensive because the medicines reach all parts of the body through the course blood. Therefore it is useful in the case of the spread of the disease. Some of chemotherapy treatment is Adriamycin, .Cytoxan, Taxol and Fluorouracil

Surgery: Is one of the methods used to treat breast cancer.. The types of surgery for the treatment of breast :cancer are **-3**

.breast sparing surgery -1

.modified radical mastectomy -2

.Partial mastectomy -3

.total mastectomy -4

-4 Hormonal therapy: Breast tumors have a strong relationship with the female hormone (estrogen). Therefore, the first methods used in the treatment of breast tumors it was hormone therapy, because it is prevents secrete or their impact on the breast. Some of Hormonal therapy treatment is Tamoxifen, Arimidex, Femara and Aromasin.

2-7. Some important information about breast cancer:

- The discovery of breast cancer early and treated early often leads to a full recovery.
- breast cancer detection late means outbreak in the body by a large margin, and it becomes difficult to treat.
- the discovery of the breast mass is not necessarily mean the presence of cancer, most breast tumors are benign.
- Eat less fat and Avoid obesity.
- Eat a lot of fiber foods and fruits and vegetables
- Check with your doctor when you see any symptoms of breast satisfactory
- Periodic inspection, (D. Mahmoud Shaheen and others).

CHAPTER THREE

SURVIVAL ANALYSIS

:Preface .3-1

In this chapter will be addressed to all the statistical techniques to be used in the practical side of this research. The main part of this chapter is to find a survival function, a function of risk, Cox model of relative risk and the comparison between two or more groups of data survive.

:Definition of Survival Analysis .3-2

Survival analysis is concerned with studying the time between entry to a study and a subsequent event and becomes one of the most important fields in statistics. The techniques developed in survival analysis are now applied in many fields, such as biology (survival time), engineering (failure time), medicine (treatment effects or the efficacy of drugs), quality control (lifetime of component), credit risk modeling in finance (default time of a firm) (Jianqing Fan 2007).

Survival analysis involves the modeling of time to event data. It has been a very active research field for several decades. An important contribution that stimulated the entire field was the counting process formulation given by Aalen (1975)).

In many biomedical studies, the outcome variable is a survival time, or more generally a time to an event. We will describe some of the standard tools for analyzing survival data.

:Survival Data .3-3

:Definition .3-3-1

Expression used to describe the data that measure the time until the event, and that the outcome variable is the time until the event which is known time to stay, which is a positive real variable values always. Kalbfleisch and Prentice (2002).

:Special Features of Survival Data .3-3-2

We must first consider the reasons why survival data are not amenable to standard statistical procedures used in data analysis. One reason is that data are generally not symmetrically distributed. Typically, a histogram constructed from the survival times of a group of similar individuals will tend to be positive skewed, that is the histogram will have a

longer "tail" to the right of the interval that contains the largest number of observations. As a consequence, it will not be reasonable to assume that data of this type have a normal distribution. This difficulty could be resolved by first transforming the data to give a more symmetric distribution Collett (2003). Second reason is that the main features of survival data that renders standard methods in appropriate is that survival times are frequently censored Cox and Oakes ((1984).

The survival time of an individual is said to be censored when the end point of interest has not been observed for that individual. This may be because the data from a study are to be analyzed at a point in time when some individuals are still alive. Alternatively, the survival status of an individual at the time of the analysis might not be known because that individual has been lost follow-up.

An actual survival time can also be regarded as censored when death is from a cause that is known to be unrelated to the treatment.

In each of these situations, a patient who entered a study at time dies at time , however, is unknown, either because the individual is still alive or because he or she has been lost to follow-up. if the individual was last known to be alive at time , the time is called a censored time.

This censoring occurs after the individual has been entered into a study, that is to the right of the last known survival time, and is therefore known as right censoring. The right-censored survival times is then less than the actual, but unknown survival time. Another form of censoring is left censoring, which is encountered when the actual survival time of an individual is less than that observed Collett (2003).

Yet another type of censoring is interval censoring. Here individuals are known to have experienced an event within an interval of time Collett (2003).

3-3-3. Patient Time and Study Time:

In atypical study, patients are not all recruited at exactly the same time. But accrue over a period of months or even years. After recruitment, patients are followed up until they die, or until a point in calendar time that marks the end of the study. When the data are analyzed, after recruitment some patients may be lost to follow up, while others will still be alive at the end of the study. The calendar time period in which an

individual is in the study is known as the study time Collett ((2003.

3-4. Survival Time Distribution:

Survival time is a variable which measures the time from a particular starting point (e.g. the time at which a treatment is initiated) to a certain endpoint of interest (e.g. the time until development of a tumor) Fox (2006).

3-4-1. Survival Function And Hazard Function:

In summarizing survival data there are two functions of the central interest, namely the survival function and the hazard function.

The actual survival time of an individual t, can be regarded as the value of a variable T, which can take any non-negative value. The different values that T can take have a probability distribution and we call The random variable associated with the survival time. Now suppose that the random variable T has A probability distribution with underlying probability density function of T is then give by

And represents the probability that the survival time is less than some value t. the survival function is defined to be the probability that the survival time is greater than or equal to t and so

The survival function can therefore be used to represent the probability that an individual survives from the time origin to sometime beyond t.

The hazard function is widely used to express the risk or the hazard of death at some time t, and is obtain from the probability that the individual dies at time t, conditional on he or she having survived to that time. For formal definition of the hazard function consider the probability that the random variable associated with an individual survival time T lies between t and , conditional on T being greater than or equal to t, written

This conditional probability is then expressed as a probability per unit time by dividing by the time interval to give a rate. The hazard function is then the limiting value of this quantity, as tends to zero, so that

The function is also referred to as the hazard rate, the instantaneous death rate, the intensity rate, or the force of mortality.

From equation (3.2) is the appropriate probability that an individual dies in the interval , conditional on that person having survived to time t. for example if the survival times is measured in days, is the approximate probability that an individual who is alive on day t, dies in the following day. For this reason, the hazard function is often simply interpreted as the risk of death at time t.

For the definition of the hazard function in equation (3.2) we can obtain some useful relationships between the survival and hazard functions.

According to a standard result from probability theory, the probability of an event A, conditional on the occurrence of an event B, is given by

where is the probability of joint occurrence of A and B. Using this result, the conditional probability is the definition of the hazard function in equation (3.2) is:

which is equal to

Where is the distribution function of T. then

Now

Is the definition of the derivative of with respect to t, which is , and so

It then follows that

and so

Where

The function features widely in survival analysis, and is called the integrated of cumulative hazard. From equation (3.5), the cumulative hazard can be obtained from the survival function since

In the analysis of survival data, the survival function and hazard function are estimated from the observed survival times Collett (2003).

3-5. Estimate of The Survival Function:

An initial step in the analysis of a set of survival data is to present numerical or graphical summaries of the survival times for individuals in a particular group. Survival data are conveniently summarised through estimates of the survival function and hazard function. The methods for estimating these functions from a single sample of survival data are said to be non-parametric or distribution-free, since they do not require specific assumptions to be made about the underlying distribution of the survival times. Once the estimated survival function or has been found the median and the other percentiles of the distribution of survival times can be estimated Collett (2003).

3-5-1. The Kaplan-Meier Estimate of The Survival Function:

Kaplan-Meier estimate of survivor and hazard functions Given n individuals with observed survival times, some of the observations may be censored and there may also be more than one individual who fails at the same observed time Therneau and Grambsch (2000). We suppose that there are individuals with observed survival times some of these observation may be right-censored, and there may also be more than one individual with the same observed survival times. We therefore suppose that there are r death times amongst the individuals. Where . After arranging these death times in ascending order the is denoted for , and so the r ordered death times are .

The number of individuals who are alive just before time , including those who are about to die at this time, will be denoted and will denote the number who die at this time Collett (2003).

We count the total number of individuals alive at the start of the interval and the number of individuals who died (di) in the time interval. The Kaplan-Meier estimate of the survival function is given by

:Standard Error of Kaplan-Meier Estimate .3-5-2

The Kaplan-Meier estimate of the survival function for any value of t in the interval from t(k) to t(k+1) can be written as for where is the estimated probability that an individual survives through ,the time interval that begins at t(j) , , taking logarithms

and so the variance of log is given by

Now the number of individuals who survive through the interval beginning at t(j) can be assumed to have a binomial distribution with parameters and , where is the true .probability of survival through that interval

The observed number who survive is , and using the result that the variance of a binomial random variable with parameters n, p is given by

since

the variance of is that is .The variance of may then be estimated by

In order to obtain the variance of , we make use of a general result for the approximate variance of a function of a random .variable

According to this result the variance of a function of the random variable x is given by

This is known as the Taylor series approximation to the variance of a function of a random. Using equation (.11) the approximate estimated variance of is , which on substitution for ,reduces to

(From equation (.9

and a further application of the result in equation (.11) gives

so that

finally the standard error of the Kaplan-Meier estimate of the survivor function, defined to be the square root of the estimated variance of the estimate is given by

.for . This result is known as Greenwood's formula
If there are no censored survival times , and expression (.12) becomes , Now

Which can be written as

. Since for

:Confidence Intervals for The Survival Function .3-5-3

Once of standard error of an estimate of the survival function has been calculated a confidence interval for the corresponding value of the survival function at a given time t .can be found

A confidence interval for the true value of the survival function at a given time t is obtained by assuming that the estimated value of the survival function at t is normally distributed with mean and estimated variance given by equation (.14). The interval computed from percentage points of the standard normal distribution. Thus, if Z is a random variable that has a standard normal distribution, the upper (one – sided) -point, or the (two sided) -point, of this distribution is that value which is such that . This probability is the area under the standard normal curve to the right of , as illustrated in figure (.1) for example the two-sided 5% and 1% points of the standard .normal distribution and , are 1.96 and 2.58, respectively

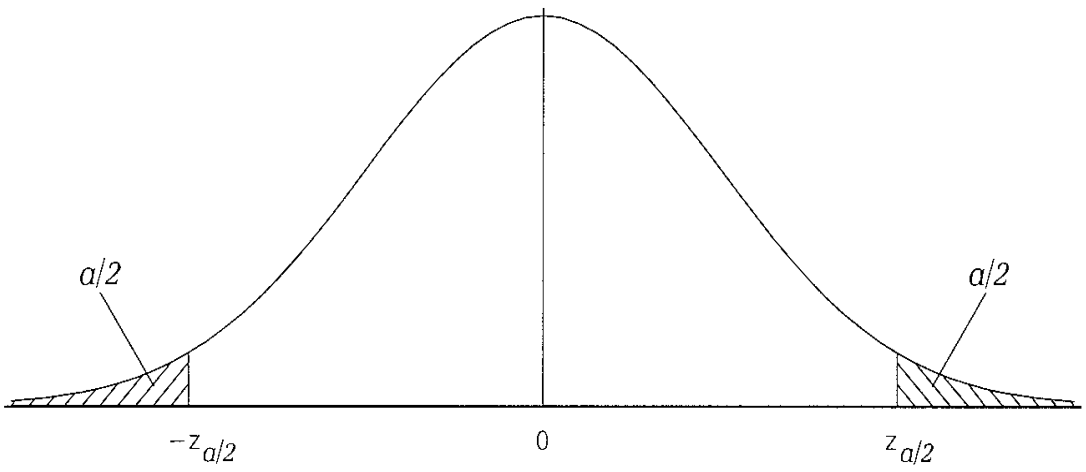


Figure (3.1): Upper and Lower -points of the standard normal distribution.

A confidence interval for S(t) for a given value of t is the interval from to is found from equation (3.15).

3-6 The Median and Percentiles of Survival Times:

3-6-1 Estimating The Median and Percentiles of Survival Times:

Since the distribution of survival times tends to be positively skew, the median is the preferred summary measure of the location of the distribution. Once the survival function has been estimated it is straight forward to obtain an estimate of the median survival times. This is the time beyond which 50% of the individuals in the population under study are expected to survive, and is given by that value which is such that . Because the non-parametric estimates of are step- functions, it will not usually be possible to realise an estimated survival time that makes the survival function exactly equal to 0.5. instead the estimated median survival time, , is defined to be the smallest observed survival times for which the value of the estimated survival function is less than 0.5.In mathematical terms.

Where t(j) is the jth ordered time, . In the particular case where the estimated survival function is exactly equal to 0.5 for values of t in the interval from to , the median is taken to be the half-way point in this interval, that is . In the situation where there are no censored survival times, the estimated median survival time will be the smallest time beyond which 50% of the individuals in the sample survive.

A similar procedure to that described above can be used to estimate other percentiles of the distribution of survival times. The percentile of the distribution of survival times is defined to be the value which is such that . In terms of the survival function is such that .

3-6-2 Confidence Interval for The Median and Percentiles:

Approximate confidence intervals for the median and percentiles of a distribution of survival times can be found once

the variance of the estimated percentile has been obtained. An expression for the approximate variance of a percentile can be derived from a direct application of the general result for the variance of a function of a random variable in equation (3.11). using this result

where is the percentile of the distribution and is the Kaplan-,Meier estimate of the survival function at . Now

An estimate of the probability density function of the survival times at , and on rearranging equation (3.16), we get

the standard error of , the estimated percentile, is therefore given by

Once the standard error of the estimated percentile has been found a confidence interval for has limits of

Where is the upper (one-sided) -point of the standard normal . distribution

:Estimating of The hazard function .3-7

A single sample of survival data may also be summarized through the hazard function, with shows the dependence of the .instantaneous risk of death on time

:Kaplan –Meier of Estimate The hazard function .3-7-1

A natural way of estimating the hazard function for unground survival data is to take ratio of the number of death at a given death time to the number of individuals at risk at that time. If the hazard function is assumed to be constant between successive death time, the hazard per unit time can be found by further dividing by the time interval. Thus if there are deaths at the death time, , and at risk at time , the hazard function in the interval from to can be estimated by

. For where
Notice that is not possible to use equation (3.18) to estimate the hazard in the interval that begins at the final death time, .since this interval is open-ended

The estimate in equation (3.18) is referred to as a Kaplan-Meier type estimate, because the estimated survival function derived from it is the Kaplan-Meier estimate.
To show this, note that since , is an estimate of the risk of death per unit time in the interval, the probability of death in that interval is , that is .Hence an estimate of the corresponding survival probability in that interval is and the estimated survival function is as given by equation (3.8).
The approximate standard error of can be found from the variance of , which may be assumed to have a binomial distribution with parameters and , where is the probability of the death in the interval of length t. Consequently , and estimating by gives

However, when is small confidence intervals constructed using this standard error will be too wide to be of practical use.

3-7-2 Estimating The cumulative hazard function:

The cumulative hazard function is important in the identification of models for survival data. The cumulative hazard at time t, H(t) was defined in equation (3.6) to be the integral of the hazard function, but is more conveniently found using equation (3.7). According to this result, , and so if is the Kaplan-Meier estimate of survival function. is an appropriate estimate of the cumulative hazard to time t.
Now using equation (3.8)

. For and , are r ordered death times with
An estimate of the cumulative hazard function also leads to an estimate of the corresponding hazard function, since the differences between adjacent values of the estimated cumulative hazard function provide estimates of the underlying hazard after dividing by the time interval.

3-8 Comparison of two or more groups of survival data:

The simplest way of comparing the survival times obtained from two or more groups of individuals is to plot the corresponding estimates of the two or more survival functions on the same axes.

3-8-1 Log-rank test for comparison of two groups of survival data:

In order to construct the long-rank test, we begin by considering separately each death time in two groups of survival data. These groups will be labeled Group I and, Group II. Suppose that there are r distinct death times, , across the two groups and that at time t(j), individuals in Group I and individuals in Group II die for . Unless two individuals in a group have the same recorded death time, the value of and will either be zero or unity. Suppose further that there are individuals at risk of death in Group I just before time t(j), and that there are at risk in Group II. Consequently at time t(j), there are deaths in total out of individuals at risk. We can therefore regard as a random variable, which can take any value in the range from zero to the minimum of and . in fact has hypergeometric distribution, according to which probability that the random variable associated with the number of deaths in the Group I takes the value is

represents the number of different ways in which times can be chosen from times and is read as . it is given by

The mean of the hypergeometric random variable is given by

So that is the expected number of individuals who die at time t(j) in Group I.
The most straight forward way of doing this is to sum the differences over the total number of death time r in the two groups.

Notice that is which the difference between the total observed and expected numbers of death in Group I. this statistic will have zero mean since . The variance of is simply the sum of the variances of the . The variance of is given by

So that the variance of is

it can be shown that has approximate normal distribution when the number of death times is not too small. It then follows that has a normal distribution with zero mean and unit variance .

The square of a standard normal random variable has a chi square distribution on one degree of freedom, denote so we have that

3-8-2. Log-rank test for comparison of three or more groups of survival data:

The long-rank test can be extended to enable three or more groups of survival data to be compared. U statistic for comparing the observed numbers of death in groups 1,2, …, g-1 with expected values. And is variance-covariance matrix.

In order to test the null hypothesis of no group differences we make use of the result the test statistic has chi-square distribution on (g-1) degrees of freedom.

3-9. Cox regression model for proportional hazards:

3-9-1. Fitting the proportional hazards model:

Given a set of covariates x and a corresponding set of coefficient β, the hazard function in the Cox model is:

The component is called “the baseline hazard function” and does not depend from the covariates x, while the exponential part of equation (3.26) is a function of the covariates x, but does not depend from the time t. These assumptions indicate that the Cox model is a PH model, as it is easily verifiable for one covariate , taking value 0 for untreated (or unexposed) subjects and 1 for the treated (or exposed) ones:

From equation (3.27) is evident that represents a measure of the association between the treatment or the exposure and the probability (i.e., the risk) of developing the outcome under study. Furthermore, in the presence of other covariates, an estimate of the HR, adjusted for the effect of such covariates, is obtained by exponentiating an estimate of , Then, one of the main advantages of the application of a regression technique, like the Cox regression model, in comparison to stratified analysis (like the MH approach), is the possibility to adjust for the effect of one or more confounders and to estimate at the same time their effect. Moreover, the Cox model allows for the introduction of continuous variables. Finally, stratified analysis in general assumes an independent effect of the considered covariates, while a regression model allows the checking for interaction between two or more predictors that can be performed introducing an interaction term among the predictors. For example, in the following equation:

represents the interaction term, obtained as the product by and , and is the corresponding coefficient, whose value will approach 0 in the case of no interaction. From equations (3.26), the Cox regression model can be expressed in terms of survival probabilities as follows:

To estimate β coefficients, Cox proposed a new method, based on “the partial likelihood” function (Cox, 1972; Cox, 1975):

where m represents the number of not censored times, and indicates that the summation is performed over all subjects in the risk set at (Hosmer and Lemeshow, 1999); β coefficients are obtained in the correspondence of the maximum values of equation (30) (maximum partial likelihood estimates) by applying mathematical procedures similar to that used in the framework of Generalized Linear Models (GLM, Dobson, 2002). Equation (30) may be transformed in a partial log-likelihood function, whose derivative respect to each coefficient βk is called a “Score” function (Hosmer and Lemeshow, 1999):

Where

Another equivalent formula is given in equation (3.31), where the sum is performed over all N observed times (i.e., including the censored ones), and the symbol “k” indicates that k coefficients, corresponding to k covariates, may be introduced into the model:

Equation (3.32) is a little more complicated than equation (3.31), but it is useful to define the so-called “Schoenfeld residuals” that are largely employed to assess the PH assumption violation (Hosmer and Lemeshow, 1999), as illustrated in a further paragraph. However, the (partial) Score function is mainly applied to estimate the β coefficients that are obtained in the correspondence of 0 values of equation (31) or, equivalently, of equation (3.32). The corresponding variance estimate may be obtained from:

where I (still similarly to the GLM approach, Dobson, 2002) is called “the observed information matrix”, and is obtained by

the second derivative of the log partial likelihood (Hosmer and Lemeshow, 1999). Statistical inference may be made using the same test applied to GLM, i.e. (partial) likelihood ratio test, score test and Wald test (Dobson, 2002), using the partial likelihood instead of the likelihood function.

3-9-2. The interpretation of estimated parameters:

The proportional hazard model can be used when the primary goal of the analysis is to estimate the effect of study variables on survival covariate . from equation (3.26), the hazard function
the interpretation of the coefficients is the difference in the log hazard corresponding to a one unit change in the covariate.

3-9-3. The hypothesis test of estimated parameters:

To test the hypothesis that a covariate has no effect we use Wald test. Is a ratio of estimated coefficient to the standard error of the estimate, and has a standard normal distribution. The hypothesis is

the formula of test written as

A confidence intervals for the estimated coefficient is obtained from the following formula:

CHAPTER FOUR

APPLICATION

:Preface .4-1

In this chapter we will use statistical methods mentioned in chapter two on the research data to obtain the required results of the study.

:Data .4-2

Data were collected from the National Center for radiotherapy and nuclear medicine in Khartoum, in the period from 2009 to 2011.

The study variables included age, date of diagnosis of the disease and even death or the date of last follow-up per weeks, education level, marital, stage, radiotherapy (given or not given), chemotherapy (given or not given), surgical (yes or no), hormonal (given or not given).

4-3. Descriptive analysis of the variables of the study:

Table (4-1) Age groups:

Not given chemotherapy		Given chemotherapy		Age groups
Percentage%	Frequency	Percentage%	Frequency	
9.3%	5	18.2%	12	lowest through 35
38.9%	21	50%	33	through 50 36
51.8%	28	31.8%	21	through 51 highest
100%	54	100%	66	**Total**

Source: prepared by the researcher by using SPSS,2014

Figure (4-1): Frequency distribution of age groups:

Source: prepared by the researcher by using Excel,2014

Seen from the table (4-1) and (Figure 4-1) that of those who given chemotherapy 33 individual by 50% were in the age group (36-50) at diagnosis, followed by age group (51 through highest) 21 individual by 31.8%, followed by that age group (lowest through 35) 12 individual by 18.2%.

While those who did not given chemotherapy 28 individual by 51.8% were in the age group (51 through highest), followed by age group (36-50) 21 individual by 38.9%, followed by that age group (lowest through 35) 5 individual by 9.3%.

Table (4-2) Educational level:

Not given chemotherapy		Given chemotherapy		Educational level
Percentage%	Frequency	Percentage%	Frequency	

42.6%	23	24.2%	16	Illiterate
31.5%	17	37.9%	25	Primary
14.8%	8	21.2%	14	High school
11.1%	6	16.7%	11	University
100%	54	100%	66	**Total**

Source: prepared by the researcher by using SPSS,2014

:Figure (4-2): Frequency distribution of education level

Source: prepared by the researcher by using Excel,2014

Seen from the table (4-2) and Figure (4-2) that of those who given chemotherapy 25 individual by 37.9% were in primary level of education, followed by (illiterate) 16 individual by 24.2%, followed (high school) 12 individual by 18.2%, followed by (university) 12 individual by 18.2%.

While those who did not given chemotherapy, 23 individual by (42.6%) were illiterate, followed primary 17 individual by 31.5%, followed high school 8 individuals by 14.8%, followed by (university) 6 individuals by 11.1%.

:Table (4-3) Marital

Not given chemotherapy		**Given chemotherapy**		**Marital**
Percentage %	**Frequency**	**Percentage %**	**Frequency**	
3.7%	2	9.1%	6	Single
90.7%	49	87.9%	58	Married
3.7%	2	3.0%	2	Divorced
1.9%	1	0.0%	0	Widowed
100%	54	100%	66	**Total**

Source: prepared by the researcher by using SPSS,2014

:Figure (4-3): Frequency distribution of marital

Source: prepared by the researcher by using Excel,2014

Seen from the table (4-3) and Figure (4-3) that of those who given chemotherapy 58 individual by 87.9% were married, followed by (single) 6 individuals by 9.1%, followed (divorced) 2 .individuals by 3.0%

While those who did not given chemotherapy, 49 individual by 90.7% were married, followed both (single and divorced) 2 .individuals by 3.7%, followed widowed one individual by 1.9%

:Table (4-4) Stages

Not given chemotherapy		**Given chemotherapy**		**Stages**
Percentag %e	**Frequency**	**Percentag %e**	**Frequency**	
5.6%	3	6.1%	4	Stage 0
9.3%	5	12.1%	8	Stage I
20.4%	11	27.3%	18	Stage II
53.7%	29	34.6%	23	Stage III
11.1%	6	19.7%	13	Stage IV
100%	54	100%	66	**Total**

Source: prepared by the researcher by using SPSS,2014

:Figure (4-4): Frequency distribution of stage

Source: prepared by the researcher by using Excel,2014

Seen from the table (4-4) and Figure (4-4) that of those who given chemotherapy 23 individual by 34.6% were in stage III, followed by (stage II) 18 individual by 27.3%, followed (stage IV) 13 individual by 19.7%, followed (stage I) 8 individual by 12.1%, .followed (stage 0) 4 individuals by 6.1%

While those who did not given chemotherapy, 29 individual by 53.7% were in stage III, followed by (stage II) 11 individual by 20.4%, followed (stage IV) 6 individuals by 11.1%, followed

(stage I) 5 individuals by 9.3%, followed (stage 0) 3 individuals by 5.6%.

:Table (4-5) : Status

Not given chemotherapy		Given chemotherapy		Status
Percentage%	Frequency	Percentage%	Frequency	
18.5%	10	56.1%	37	Censored
81.5%	44	43.9%	29	Relapsed
100%	54	100%	66	**Total**

Source: prepared by the researcher by using SPSS,2014

:Figure (4-5): Frequency distribution of status

Source: prepared by the researcher by using Excel,2014

Seen from the table (4-5) and Figure (4-5) that of those who given chemotherapy 37 individual by (56.1%) were censored, followed by relapsed 29 individual by (43.9%).
While those who did not given chemotherapy, 44 individual by (81.5%) were relapsed, followed by censored 10 individuals by (18.5%)).

:Table (4-6): Radiotherapy

Not given chemotherapy		Given chemotherapy		Radiotherapy
Percentage%	Frequency	Percentage%	Frequency	
77.8%	42	78.8%	52	Given
22.2%	12	21.2%	14	Not given
100%	54	100%	66	**Total**

Source: prepared by the researcher by SPSS,2014

:Figure (4-6): Frequency distribution of radiotherapy

Source: prepared by the researcher by using Excel,2014

Seen from the table (4-6) and Figure (4-6) that of those who given chemotherapy 52 individual by (78.8%) were given radiotherapy, followed by not given radiotherapy 14 individual by (21.2%).

While those who did not given chemotherapy, 42 individual by (77.8%) were given radiotherapy, followed by not given radiotherapy 12 individual by (22.2%).

:Table (4-7): Surgical

Not given chemotherapy		Given chemotherapy		Surgical
Percentage%	Frequency	Percentage%	Frequency	
9.3%	5	16.7%	11	Yes
90.7%	49	83.3%	55	No

100%	54	100%	66	**Total**

Source: prepared by the researcher by SPSS,2014

Figure (4-7): Frequency distribution of surgical

Source: prepared by the researcher by using Excel,2014

Seen from the table (4-7) and Figure (4-7) that of those who given chemotherapy 55 individual by (83.3%) did not undergo surgery, followed by whose underwent surgery 11 individual by ((16.7%.

While those who did not given chemotherapy, 49 individual by (90.7%) did not undergo surgery, followed by whose underwent surgery 5 individuals by (9.3%).

:Table (4-8): Hormonal

Not given chemotherapy		**Given chemotherapy**		**Hormonal**
Percentag %e	**Frequency**	**Percentag %e**	**Frequency**	
33.3%	18	30.3%	20	Yes
66.7%	36	69.7%	46	No
100%	54	100%	66	**Total**

Source: prepared by the researcher by using SPSS,2014

:Figure (4-8): Frequency distribution of hormonal

Source: prepared by the researcher by using Excel,2014

Seen from the table (4-8) and Figure (4-8) that of those who given chemotherapy 46 individual by (69.7%) were not given Hormonal therapy, followed by whose given Hormonal therapy (20 individual by (30.3%.

While those who did not given chemotherapy, 36 individual by (66.7%) were not given Hormonal therapy, followed by whose given Hormonal therapy 18 individual by (33.3%).

Kaplan-Meier to estimate survival function , .4-4 standard error and confidence intervals at the 5% level of significance:

Table (4-9): Kaplan-Meier to estimate survival function , standard error and confidence intervals at the 5% level of significance for patients who were given chemotherapy:

confidence intervals		standard error	Survival function	Number of censors	Number of deaths	Number of surviving	Time
Upper	Lower						
.	.	.	1.0000	1	0	66	3
0.9849	0.8637	0.0260	0.9538	0	3	65	6
0.9764	0.8443	0.0298	0.9385	0	1	62	8
0.9672	0.8250	0.0331	0.9231	3	1	61	9
0.9672	0.8250	0.0331	0.9231	2	0	57	10
0.9672	0.8250	0.0331	0.9231	1	0	55	12
0.9567	0.8024	0.0366	0.9060	1	1	54	13
0.9335	0.7583	0.0427	0.8711	0	2	52	15
0.9335	0.7583	0.0427	0.8711	1	0	50	16
0.9211	0.7365	0.0454	0.8534	0	1	49	19
0.9211	0.7365	0.0454	0.8534	1	0	48	20
0.9080	0.7145	0.0479	0.8352	0	1	47	23
0.8946	0.6931	0.0502	0.8170	0	1	46	24
0.8808	0.6721	0.0522	0.7989	0	1	45	26

0.8524	0.6312	0.0558	0.7626	0	2	44	30
0.8378	0.6112	0.0574	0.7444	1	1	42	32
0.8378	0.6112	0.0574	0.7444	1	0	40	36
0.8378	0.6112	0.0574	0.7444	1	0	39	38
0.8221	0.5893	0.0591	0.7248	1	1	38	44
0.8058	0.5670	0.0608	0.7047	0	1	36	46
0.7553	0.5024	0.0648	0.6443	0	3	35	50
0.7553	0.5024	0.0648	0.6443	1	0	32	54
0.7375	0.4806	0.0660	0.6235	1	1	31	60
0.7190	0.4581	0.0671	0.6020	0	1	29	62
0.7190	0.4581	0.0671	0.6020	1	0	28	72
0.6997	0.4350	0.0682	0.5797	2	1	27	80
0.6578	0.3851	0.0706	0.5314	1	2	24	86
0.6129	0.3343	0.0724	0.4808	0	2	21	89
0.6129	0.3343	0.0724	0.4808	1	0	19	94
0.6129	0.3343	0.0724	0.4808	3	0	18	96
0.6129	0.3343	0.0724	0.4808	2	0	15	115
0.6129	0.3343	0.0724	0.4808	2	0	13	120
0.6129	0.3343	0.0724	0.4808	1	0	11	122
0.6129	0.3343	0.0724	0.4808	1	0	10	128
0.5786	0.2669	0.0817	0.4274	0	1	9	133
0.5390	0.2090	0.0872	0.3740	0	1	8	138
0.5390	0.2090	0.0872	0.3740	3	0	7	140
0.5390	0.2090	0.0872	0.3740	4	0	4	144

Source: prepared by the researcher by using STATA,2014

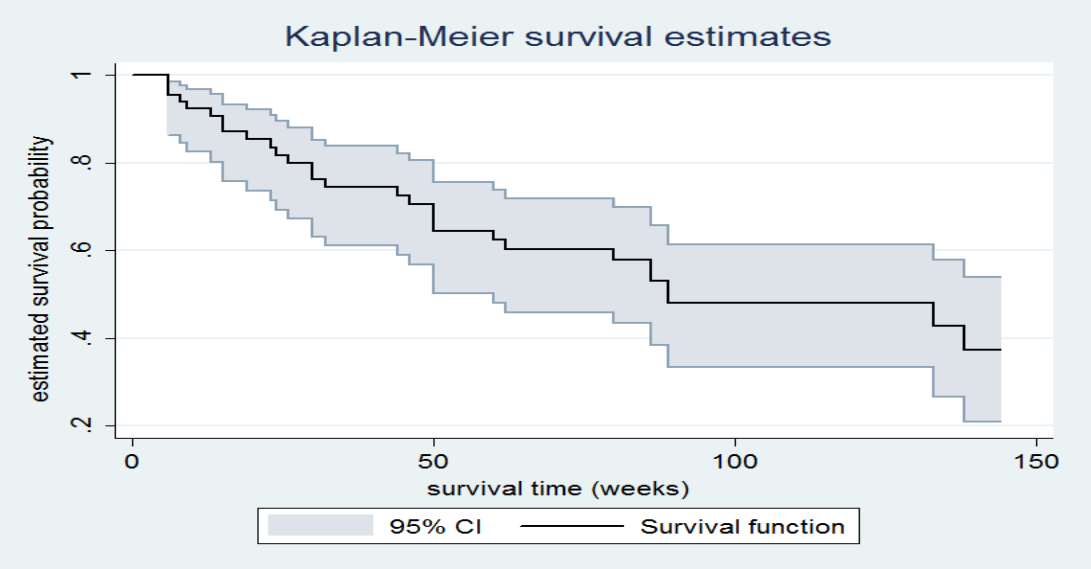
Seen from table (4-9) that the first survival time observed is (3) weeks, and there are (66) individuals at risk, and the estimated survival function at any point in the interval [0-3) .equal to one. and there are (1) lost follow-up

The second observed survival time is (6) weeks there are (65) individuals at risk, (3) individuals are deaths, the value of the estimated survival function is (0.9538), with standard error (0.0260) and confidence intervals (0.8637 – 0.9849) at 5% significance level. the value of the function remains at this .value until the time of death observed below

the third observed survival time is (8) weeks, and the number of individuals at risk is (62), there is (1) death. the value of the estimated survival function is (0.9385), with standard error (0.0296) and confidence intervals (0.8443 – 0.9764) at 5% significance level. the value of the function remains at this .value until the time of death observed below

This process continues until the last time of death (144) weeks, where there are (4) individuals at risk and died. the value of the estimated survival function is (0.3740), and .remain at this value

Figure (4-9): Kaplan-Meier to estimate survival function, confidence intervals at the 5% level of significance for patients who were given chemotherapy:



Source: prepared by the researcher by using STATA,2014

We note from figure (4-9): We note that the survival curve remained constant until observe time of death (3) weeks, and then became a function of decreasing.

Table (4-10): Kaplan-Meier to estimate survival function , standard error and confidence intervals at the 5% level of significance for patients who were not given chemotherapy:

confidence intervals		standard error	Survival function	Number of censors	Number of deaths	Number of surviving	Time
Upper	Lower						
0.9974	0.8757	0.0183	0.9815	0	1	54	1
0.9906	0.8599	0.0257	0.9630	0	1	53	2
0.9817	0.8376	0.0312	0.9444	0	1	52	3
0.9485	0.7693	0.0428	0.8889	0	3	51	6
0.9485	0.7693	0.0428	0.8889	1	0	48	7
0.9485	0.7693	0.0428	0.8889	1	0	47	10
0.9356	0.7457	0.0460	0.8696	1	1	46	13
0.9078	0.6985	0.0517	0.8300	2	2	44	15
0.8927	0.6740	0.0544	0.8093	0	1	40	19

0.8612	0.6267	0.0590	0.7678	0	2	39	20
0.8449	0.6037	0.0610	0.7470	0	1	37	23
0.8283	0.5811	0.0627	0.7263	0	1	36	26
0.8113	0.5589	0.0642	0.7055	1	1	35	30
0.7758	0.5136	0.0671	0.6628	0	2	33	32
0.7575	0.4915	0.0682	0.6414	0	1	31	36
0.7013	0.4271	0.0708	0.5773	1	3	30	38
0.6410	0.3622	0.0723	0.5106	0	3	26	43
0.6203	0.3412	0.0725	0.4884	0	1	23	44
0.5994	0.3206	0.0725	0.4662	0	1	22	52
0.5783	0.3002	0.0724	0.4440	0	1	21	55
0.5568	0.2802	0.0721	0.4218	0	1	20	62
0.5568	0.2802	0.0721	0.4218	1	0	19	64
0.5342	0.2590	0.0718	0.3984	0	1	18	69
0.5112	0.2383	0.0713	0.3750	0	1	17	70
0.4880	0.2180	0.0706	0.3515	0	1	16	74
0.4643	0.1981	0.0697	0.3281	0	1	15	80
0.4403	0.1787	0.0685	0.3047	0	1	14	86
0.3911	0.1412	0.0655	0.2578	0	2	13	88
0.3401	0.1059	0.0614	0.2109	0	2	11	89
0.3401	0.1059	0.0614	0.2109	1	0	9	90
0.3401	0.1059	0.0614	0.2109	1	0	8	98
0.3096	0.0825	0.0596	0.1808	0	1	7	104
0.2776	0.0611	0.0568	0.1507	0	1	6	112
0.2441	0.0418	0.0528	0.1205	0	1	5	116
0.1709	0.0115	0.0401	0.0603	0	2	4	120
0.1307	0.0024	0.0292	0.0301	0	1	2	122
.	.	.	0.0000	0	1	1	138

Source: prepared by the researcher by using STATA,2014

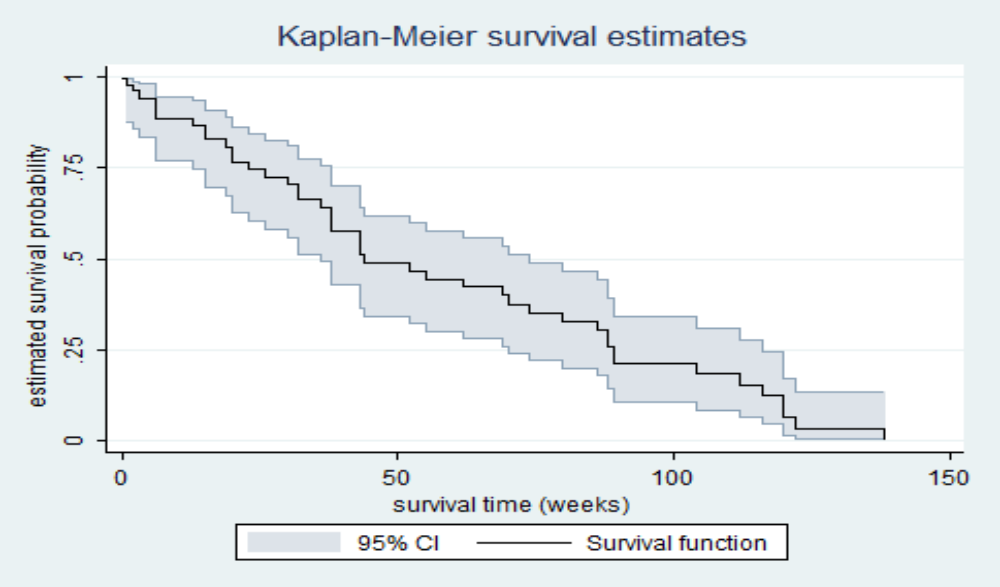
Seen from table (4-10) that the first survival time observed is (1) week, and there are (54) individuals at risk, (1) individual is deaths. the value of the estimated survival function is (0.9815), with standard error (0.0183) and confidence intervals (0.8757 – 0.9974) at 5% significance level. the value of the function remains at this value until the time of death observed below.

The second observed survival time is (2) weeks there are (53) individuals at risk, (1) individual is deaths, the value of the estimated survival function is (0.9630), with standard error (0.0257) and confidence intervals (0.8599 – 0.9906) at 5% significance level. the value of the function remains at this value until the time of death observed below.

the third observed survival time is (1) weeks, and the number of individuals at risk is (52), there is (1) death. the value of the estimated survival function is (0.9444), with standard error

(0.0312) and confidence intervals (0.8376 – 0.9817) at 5% significance level. the value of the function remains at this .value until the time of death observed below
This process continues until the last time of death (138) weeks, where there are (1) individual at risk and died. the .(value of the estimated survival function is (0.000

Figure (4-10): Kaplan-Meier to estimate survival function, confidence intervals at the 5% level of significance for patients :who were not given chemotherapy



Source: prepared by the researcher by using STATA,2014

We note from figure (4-10): We note that the survival curve is .decreasing from observe time of death (1) weeks

Kaplan-Meier to estimate hazard function , standard .4-5 error and confidence intervals at the 5% level of :significance

Table (4-11): Kaplan-Meier to estimate hazard function , standard error and confidence intervals at the 5% level of :significance for patients who were given chemotherapy

confidence intervals	standard	Survival	Number of	Number of	Number of	Time

Upper	Lower	error	function	censors	deaths	surviving	
.	.	.	0.0000	1	0	66	3
0.1363	0.0151	0.0260	0.0462	0	3	65	6
0.1557	0.0236	0.0298	0.0615	0	1	62	8
0.1750	0.0328	0.0331	0.0769	3	1	61	9
0.1750	0.0328	0.0331	0.0769	2	0	57	10
0.1750	0.0328	0.0331	0.0769	1	0	55	12
0.1976	0.0433	0.0366	0.0940	1	1	54	13
0.2417	0.0665	0.0427	0.1289	0	2	52	15
0.2417	0.0665	0.0427	0.1289	1	0	50	16
0.2635	0.0789	0.0454	0.1466	0	1	49	19
0.2635	0.0789	0.0454	0.1466	1	0	48	20
0.2855	0.0920	0.0479	0.1648	0	1	47	23
0.3069	0.1054	0.0502	0.1830	0	1	46	24
0.3279	0.1192	0.0522	0.2011	0	1	45	26
0.3688	0.1476	0.0558	0.2374	0	2	44	30
0.3888	0.1622	0.0574	0.2556	1	1	42	32
0.3888	0.1622	0.0574	0.2556	1	0	40	36
0.3888	0.1622	0.0574	0.2556	1	0	39	38
0.4107	0.1779	0.0591	0.2752	1	1	38	44
0.4330	0.1942	0.0608	0.2953	0	1	36	46
0.4976	0.2447	0.0648	0.3557	0	3	35	50
0.4976	0.2447	0.0648	0.3557	1	0	32	54
0.5194	0.2625	0.0660	0.3765	1	1	31	60
0.5419	0.2810	0.0671	0.3980	0	1	29	62
0.5419	0.2810	0.0671	0.3980	1	0	28	72
0.5650	0.3003	0.0682	0.4203	2	1	27	80
0.6149	0.3422	0.0706	0.4686	1	2	24	86
0.6657	0.3871	0.0724	0.5192	0	2	21	89
0.6657	0.3871	0.0724	0.5192	1	0	19	94
0.6657	0.3871	0.0724	0.5192	3	0	18	96
0.6657	0.3871	0.0724	0.5192	2	0	15	115
0.6657	0.3871	0.0724	0.5192	2	0	13	120
0.6657	0.3871	0.0724	0.5192	1	0	11	122
0.6657	0.3871	0.0724	0.5192	1	0	10	128
0.7331	0.4214	0.0817	0.5726	0	1	9	133
0.7910	0.4610	0.0872	0.6260	0	1	8	138
0.7910	0.4610	0.0872	0.6260	3	0	7	140
0.7910	0.4610	0.0872	0.6260	4	0	4	144

Source: prepared by the researcher by using STATA,2014

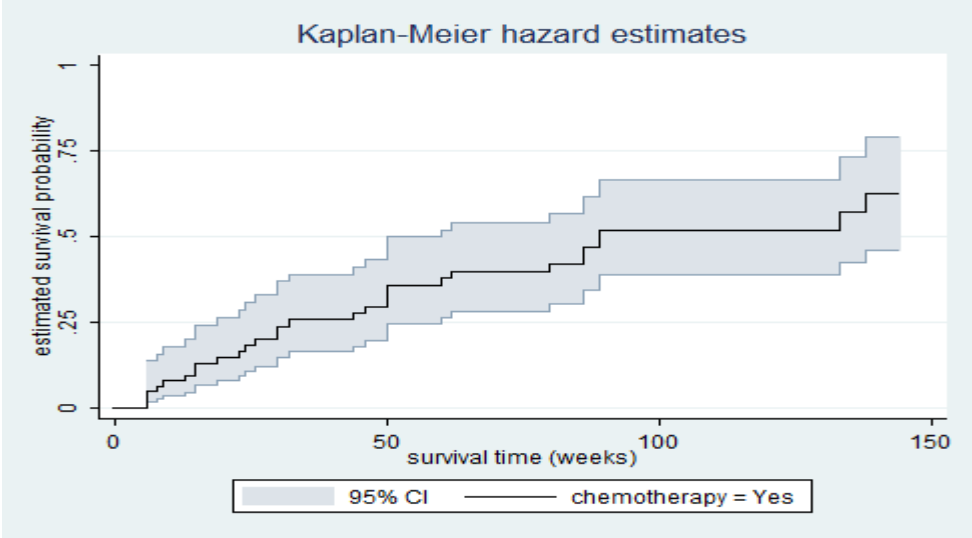
Seen from table (4-11) that the first deaths time observed is (3) weeks, and there are (66) individuals at risk, (1) individual is lost follow-up. the value of the estimated hazard function is .((0.000

The second observed deaths time is (6) weeks there are (65) individuals at risk, (3) individual is deaths, the value of the estimated hazard function is (0.0462), with standard error (0.0260) and confidence intervals (0.0151 – 0.1363) at 5% significance level. the value of the function remains at this .value until the time of death observed below

the third observed deaths time is (8) weeks, and the number of individuals at risk is (62), there is (1) death. the value of the estimated hazard function is (0.0615), with standard error (0.0298) and confidence intervals (0.0236 – 0.1557) at 5% significance level. the value of the function remains at this .value until the time of death observed below

This process continues until the last time of death (144) weeks, where there are (4) individual at risk and are lost follow-up. the value of the estimated hazard function is (0.6260), with standard error (0.0872) and confidence .intervals (0.4610 – 0.7910) at 5% significance level

Figure (4-11): Kaplan-Meier to estimate hazard function, confidence intervals at the 5% level of significance for patients :who were given chemotherapy



Source: prepared by the researcher by using STATA,2014

We note from figure (4-11): We note that the survival curve function was equal to zero until observed time (3) then .become an incremental

Table (4-12): Kaplan-Meier to estimate hazard function , standard error and confidence intervals at the 5% level of :significance for patients who were not given chemotherapy

confidence intervals		standard error	Survival function	Number of censors	Number of deaths	Number of surviving	Time
Upper	Lower						
0.1243	0.0026	0.0183	0.0185	0	1	54	1
0.1401	0.0094	0.0257	0.0370	0	1	53	2
0.1624	0.0183	0.0312	0.0556	0	1	52	3
0.2307	0.0515	0.0428	0.1111	0	3	51	6
0.2307	0.0515	0.0428	0.1111	1	0	48	7
0.2307	0.0515	0.0428	0.1111	1	0	47	10
0.2543	0.0644	0.0460	0.1304	1	1	46	13
0.3015	0.0922	0.0517	0.1700	2	2	44	15
0.3260	0.1073	0.0544	0.1907	0	1	40	19
0.3733	0.1388	0.0590	0.2322	0	2	39	20
0.3963	0.1551	0.0610	0.2530	0	1	37	23
0.4189	0.1717	0.0627	0.2737	0	1	36	26
0.4411	0.1887	0.0642	0.2945	1	1	35	30
0.4864	0.2242	0.0671	0.3372	0	2	33	32
0.5085	0.2425	0.0682	0.3586	0	1	31	36
0.5729	0.2987	0.0708	0.4227	1	3	30	38
0.6378	0.3590	0.0723	0.4894	0	3	26	43
0.6588	0.3797	0.0725	0.5116	0	1	23	44
0.6794	0.4006	0.0725	0.5338	0	1	22	52
0.6998	0.4217	0.0724	0.5560	0	1	21	55
0.7198	0.4432	0.0721	0.5782	0	1	20	62
0.7198	0.4432	0.0721	0.5782	1	0	19	64
0.7410	0.4658	0.0718	0.6016	0	1	18	69
0.7617	0.4888	0.0713	0.6250	0	1	17	70
0.7820	0.5120	0.0706	0.6485	0	1	16	74
0.8019	0.5357	0.0697	0.6719	0	1	15	80
0.8213	0.5597	0.0685	0.6953	0	1	14	86
0.8588	0.6089	0.0655	0.7422	0	2	13	88
0.8941	0.6599	0.0614	0.7891	0	2	11	89
0.8941	0.6599	0.0614	0.7891	1	0	9	90
0.8941	0.6599	0.0614	0.7891	1	0	8	98
0.9175	0.6904	0.0596	0.8192	0	1	7	104
0.9389	0.7224	0.0568	0.8493	0	1	6	112
0.9582	0.7559	0.0528	0.8795	0	1	5	116
0.9885	0.8291	0.0401	0.9397	0	2	4	120
0.9976	0.8693	0.0292	0.9699	0	1	2	122
.	.	.	1.0000	0	1	1	138

Source: prepared by the researcher by using STATA,2014

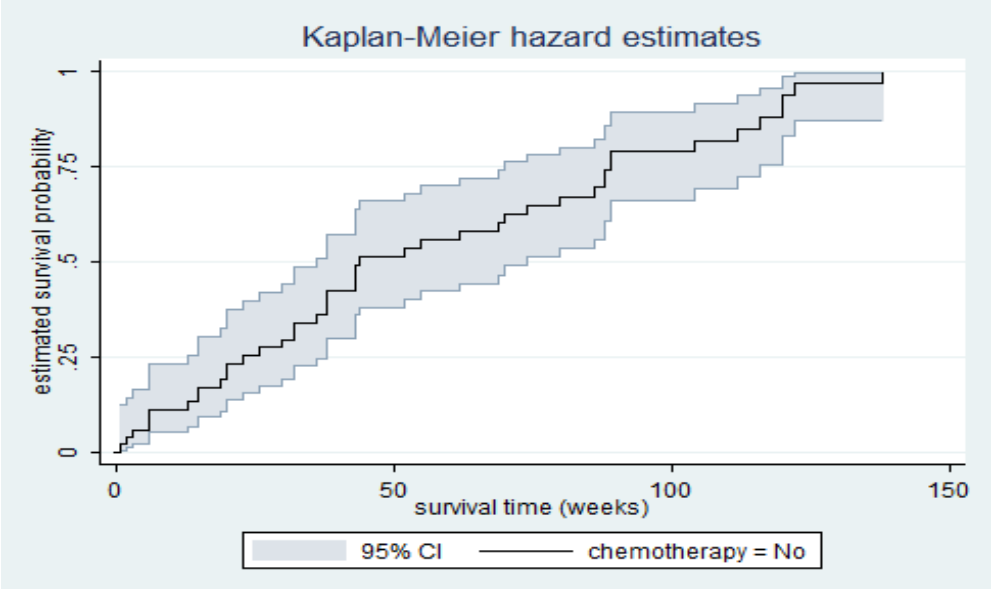
Seen from table (4-12) that the first deaths time observed is (1) week, and there are (54) individuals at risk, (1) individual is death. the value of the estimated hazard function is (0.0185), with standard error (0.0183) and confidence intervals (0.0026 – 0.1243) at 5% significance level. the value of the function remains at this value until the time of death observed below.

The second observed deaths time is (2) weeks there are (53) individuals at risk, (1) individual is deaths, the value of the estimated hazard function is (0.0370), with standard error (0.0257) and confidence intervals (0.0094 – 0.1401) at 5% significance level. the value of the function remains at this value until the time of death observed below.

the third observed deaths time is (3) weeks, and the number of individuals at risk is (52), there is (1) death. the value of the estimated hazard function is (0.0556), with standard error (0.0312) and confidence intervals (0.0183 – 0.1624) at 5% significance level. the value of the function remains at this value until the time of death observed below.

This process continues until the last time of death (138) weeks, where there are (1) individual at risk and is death. the value of the estimated hazard function is (1.000).

Figure (4-12): Kaplan-Meier to estimate hazard function, confidence intervals at the 5% level of significance for patients who were not given chemotherapy:



Source: prepared by the researcher by using STATA,2014

We note from figure (4-12): We note that the survival curve is increase.

4-6: Estimation of Median and Quartiles:

4-6-1 Estimation of median and quartiles for survival time:

Table (4-13): quartiles estimated for survival time:

Confidence 95% intervals		Standard error	Estimate	Quartile
Upper	Lower			
38	20	5.7835	30	25
88	50	12.3322	70	50
.	104	8.6091	122	75

Source: prepared by the researcher by using STATA, 2014

the first quartile of each individuals is (30) weeks, this means that 25% of the individuals will live (30) weeks, and it does not at least 20 weeks and not more than 38 weeks.

the median of each individuals is (70) weeks, this means that 50% of the individuals will live (70) weeks, and it does not at least 50 weeks and not more than 88 weeks.

the third quartile of each individuals is (122) weeks, this means that 75% of the individuals will live (122) weeks, and it does not at least 104.

4-6-2 Estimation quartiles of survival time for patients whose given chemotherapy and whose do not given chemotherapy:

4-6-2-1 Estimation first quartile of survival time for patients whose given chemotherapy and whose do not given chemotherapy:

Table (4-14): Estimation first quartile of survival time for patients whose given chemotherapy and whose do not given chemotherapy

Confidence 95%		Standard	Estimat	Number of	Chemothera

intervals		error	e	Patients	py
Upper	Lower				
60	19	8.6346	32	66	Given
38	13	5.4085	23	54	Not given
38	20	5.7835	30	120	Total

Source: prepared by the researcher by using STATA, 2014

the first quartile estimation for patients whose given chemotherapy is (32) weeks, this means that 25% of the patients whose given chemotherapy will live (32) weeks, and it does not at least 19 weeks and not more than 60 weeks.

and the first quartile estimation for patients whose did not given chemotherapy is (23) weeks, this means that 25% of the patients whose did not given chemotherapy will live (23) weeks, and it does not at least 13 weeks and not more than 38 weeks.

Estimation median of survival time for patients whose given chemotherapy and whose do not given chemotherapy: .4-6-2-2

Table (4-15): Estimation median of survival time for patients whose given chemotherapy and whose do not given chemotherapy:

Confidence 95% intervals		Standard error	Estimate	Number of Patients	Chemotherapy
Upper	Lower				
.	60	25.1766	89	66	Given
74	36	9.2495	44	54	Not given
88	50	12.3322	70	120	Total

Source: prepared by the researcher by using STATA, 2014

the median estimation for patients whose given chemotherapy is (89) weeks, this means that 50% of the patients whose given chemotherapy will live (89) weeks, and it does not at least 60 weeks.

and the median estimation for patients whose did not given chemotherapy is (44) weeks, this means that 50% of the patients whose did not given chemotherapy will live (44) weeks, and it does not at least 37 weeks and not more than 74 weeks.

Estimation third quartile of survival time for .4-6-2-3 patients whose given chemotherapy and whose do not : given chemotherapy

Table (4-16): Estimation third quartile of survival time for patients whose given chemotherapy and whose do not given :chemotherapy

Confidence 95% intervals		Standard error	Estimate	Number of Patients	Chemotherapy
Upper	Lower				
.	138	.	.	66	Given
116	70	8.9225	89	54	Not given
.	104	8.6091	122	120	Total

Source: prepared by the researcher by using STATA, 2014

the third quartile estimation for patients whose did not given chemotherapy is (89) weeks, this means that 75% of the patients whose did not given chemotherapy will live (89) weeks, .and it does not at least 70 weeks and not more than 116 weeks

:Univariate analysis .4-7

the Kaplan –Meier survival curves for different patient groups, and introduced the log-rank test to investigate differences between them. Both these methods are examples of univariate analysis; they describe the survival with respect to the factor under investigation, but necessarily ignore the impact of any others. It is more common, at least in clinical investigations, to have a situation where several (known) quantities or covariates, potentially affect patient prognosis. when investigating survival in relation to any one factor, it is often desirable to .(adjust for the impact of others Bradburn (2003

:Log-rank test for equality of survival functions .4-7-1

the log-rank test provides a P-value for the differences between the groups, it offers no estimate of the actual effect size; in other words, it offers a statistical, but not a clinical, assessment of the factor's impact. The use of a statistical model improves on these methods by allowing survival to be assessed with respect to several factors simultaneously, and in addition, offers estimates of the strength of effect for each constituent factor Bradburn ((2003.

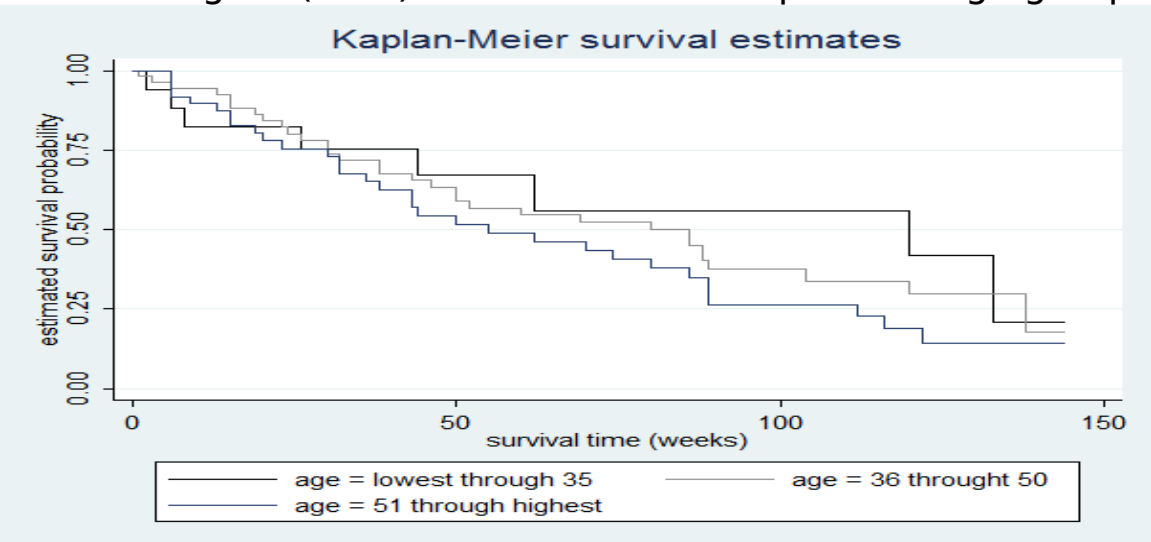
Log-rank test for equality of survival functions .4-7-1-1 :for Age groups

Table (4-17) Log-rank test for equality of estimated survival :functions for Age groups

P-value	Chi-square test	Events expected	Events observed	Age Group
0.4132	1.77	10.41	8	Lowest through 35
		35.66	33	through 50 36
		26.93	32	through 51 highest
		73.00	73	**Total**

Source: prepared by the researcher by using STATA, 2014

Figure (4-13): survival curves of patients age groups



Source: prepared by the researcher by using STATA, 2014

We note from the table (4-17) that the value of chi-square test was (1.77), and the significant value to it (P-value =

0.4132 >0.05), there was non-significant difference between .the estimated survival functions for the age group
We note from figure (4-13) there is no difference between the curves of Kaplan - Meier survival functions for age group. Thus, the probability of survival does not vary according to .age group

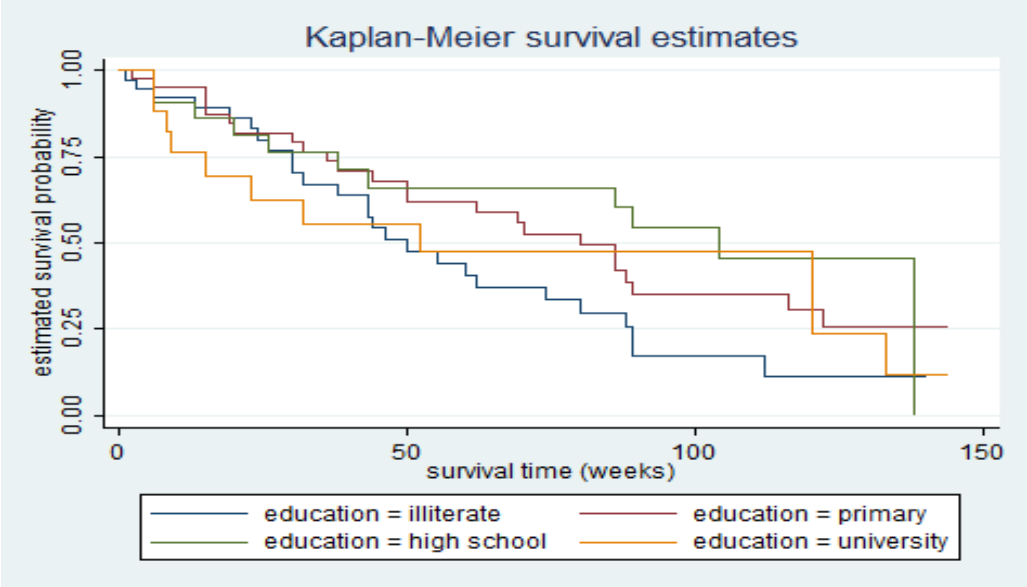
Log-rank test for equality of survival functions .4-7-1-2 :for Education levels

Table (4-18) Log-rank test for equality of estimated survival :functions for Education levels

P-value	Chi-square test	Events expected	Events observed	Education
0.1930	4.73	19.06	26	Illiterate
		28.06	24	Primary
		16.37	12	High school
		9.51	11	University
		73.00	73	**Total**

Source: prepared by the researcher by using STATA, 2014

Figure (4-14): survival curves of patients education



Source: prepared by the researcher by using STATA, 2014

We note from the table (4-18) that the value of chi-square test was (4.73), and the significant value to it (P-value = 0.1930 >0.05), there was non-significant difference between the estimated survival functions for the education level. We note from figure (4-14) there is no difference between the

curves of Kaplan - Meier survival functions for education level. Thus, the probability of survival does not vary according to education level.

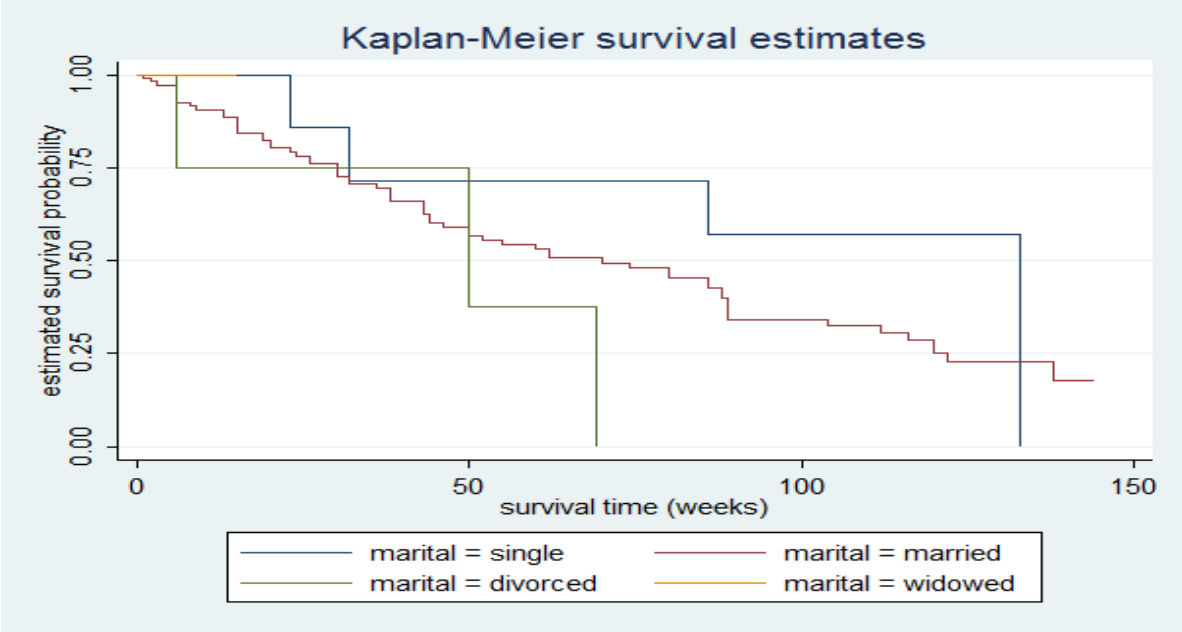
Log-rank test for equality of survival functions .4-7-1-3 :for Marital

Table (4-19) Log-rank test for equality of estimated survival functions for marital:

P-value	Chi-square test	Events expected	Events observed	Marital
		6.39	4	Single
		65.07	66	Married
0.3864	3.03	1.39	3	Divorced
		0.16	0	Widowed
		73.00	73	**Total**

Source: prepared by the researcher by using STATA, 2014

Figure (4-15): survival curves of patients marital



Source: prepared by the researcher by using STATA, 2014

We note from the table (4-19) that the value of chi-square test was (3.03), and the significant value to it (P-value = 0.3864 >0.05), there was non-significant difference between the estimated survival functions for the marital. We note from figure (4-15) there is no difference between the curves of

Kaplan - Meier survival functions for marital. Thus, the probability of survival does not vary according to marital.

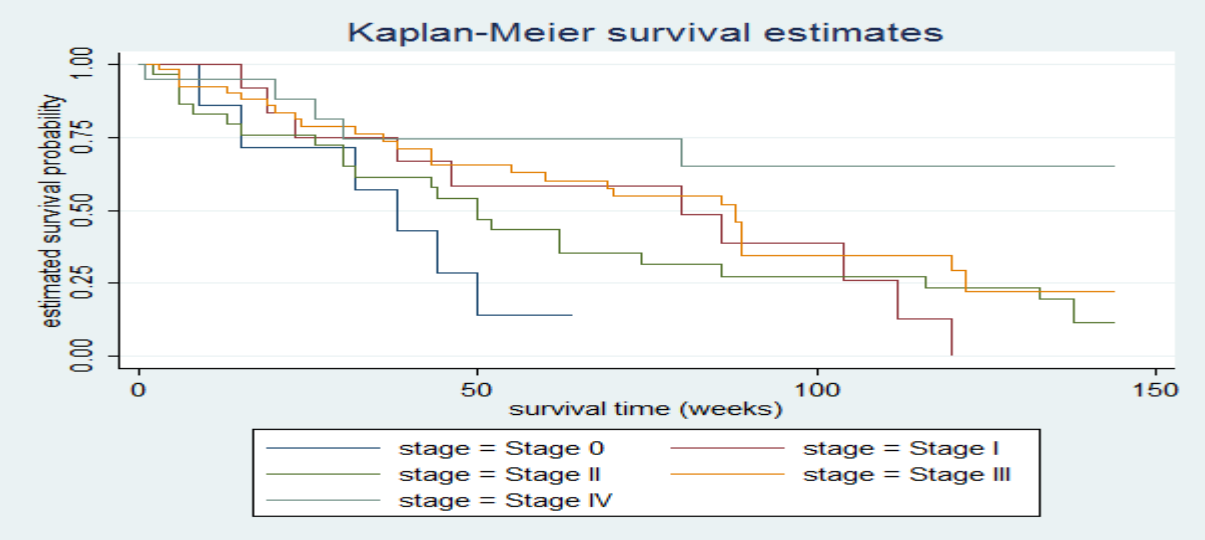
Log-rank test for equality of survival functions .4-7-1-4 :for Stages

Table (4-20) Log-rank test for equality of estimated survival :functions for stages

P-value	Chi-square test	Events expected	Events observed	Stages
0.0294	10.76	2.67	6	Stage 0
		8.26	10	Stage I
		19.02	24	Stage II
		30.87	28	Stage III
		12.18	5	Stage IV
		73.00	73	**Total**

Source: prepared by the researcher by using STATA, 2014

Figure (4-16): survival curves of patients stage



Source: prepared by the researcher by using STATA, 2014

We note from the table (4-20) that the value of chi-square test was (10.76), and the significant value to it (P-value = 0.0294 <0.05), there was significant difference between the

estimated survival functions for stage. We note from figure (4-16) there is no difference between the curves of Kaplan - Meier survival functions for stage. Thus, the probability of survival does not vary according to stage.

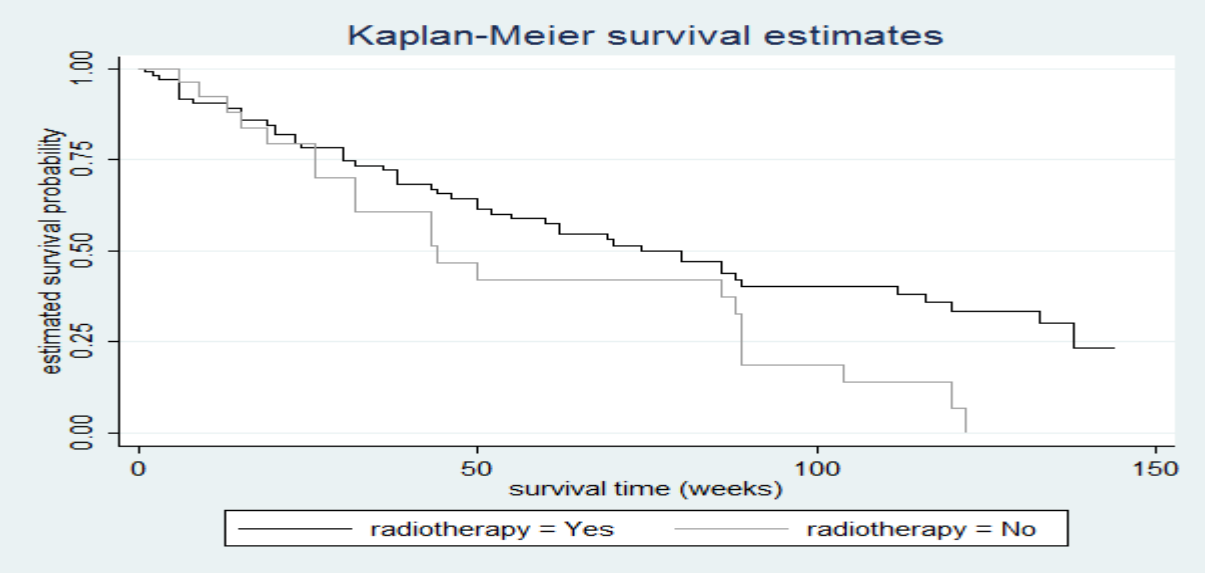
Log-rank test for equality of survival functions .4-7-1-5 for Radiotherapy:

Table (4-21) Log-rank test for equality of estimated survival functions for radiotherapy:

Radiotherapy	Events observed	Events expected	Chi-square test	P-value
(Given (yes	52	59.09		
Not given ((no	21	13.91	4.65	0.0310
Total	73	73.00		

Source: prepared by the researcher by using STATA, 2014

Figure (4-17): survival curves of patients radiotherapy



Source: prepared by the researcher by using STATA, 2014

We note from the table (4-21) that the value of chi-square test was (4.65), and the significant value to it (P-value = 0.0310 <0.05), there was significant difference between the estimated survival functions for radiotherapy. We note from figure (4-17) there is no difference between the curves of Kaplan - Meier survival functions for radiotherapy. Thus, the probability of survival does not vary according to .radiotherapy

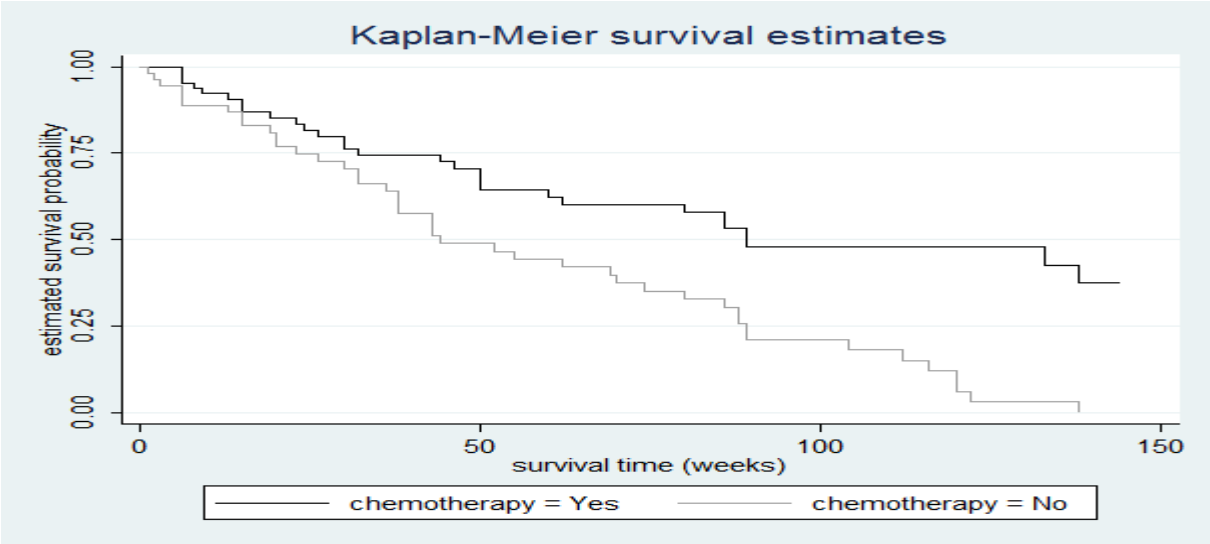
Log-rank test for equality of survival functions .4-7-1-6 :for Chemotherapy

Table (4-22) Log-rank test for equality of estimated survival :functions for chemotherapy

P-value	Chi-square test	Events expected	Events observed	Chemotherapy
0.0001	14.73	44.55	29	(Given (yes
		28.45	44	Not given ((no
		73.00	73	**Total**

Source: prepared by the researcher by using STATA, 2014

Figure (4-18): survival curves of patients chemotherapy



Source: prepared by the researcher by using STATA, 2014

We note from the table (4-22) that the value of chi-square test was (14.73), and the significant value to it (P-value = 0.0001 <0.05), there was significant difference between the estimated survival functions for chemotherapy. We note from figure (4-18) there is difference between the curves of Kaplan - Meier survival functions for chemotherapy. Thus, the probability of survival does not vary according to .chemotherapy

Log-rank test for equality of survival functions .4-7-1-7 :for Surgical

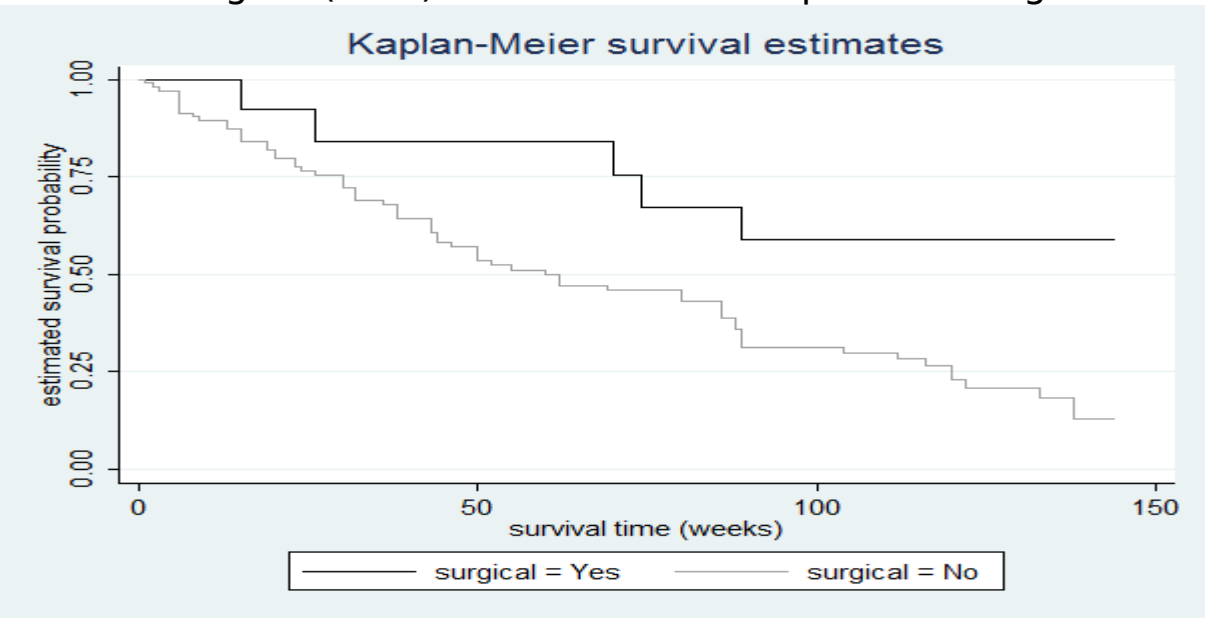
Table (4-23) Log-rank test for equality of estimated survival :functions for surgical

P-value	Chi-square test	Events expected	Events observed	Surgical
0.0285	4.80	11.72	5	Yes

	61.28	68	No
	73.00	73	**Total**

Source: prepared by the researcher by using STATA, 2014

Figure (4-19): survival curves of patients surgical



Source: prepared by the researcher by using STATA, 2014

We note from the table (4-23) that the value of chi-square test was (4.80), and the significant value to it (P-value = 0.0285 <0.05), there was significant difference between the estimated survival functions for surgical. We note from figure (4-19) there is difference between the curves of Kaplan - Meier survival functions for surgical. Thus, the probability of .survival vary according to surgical

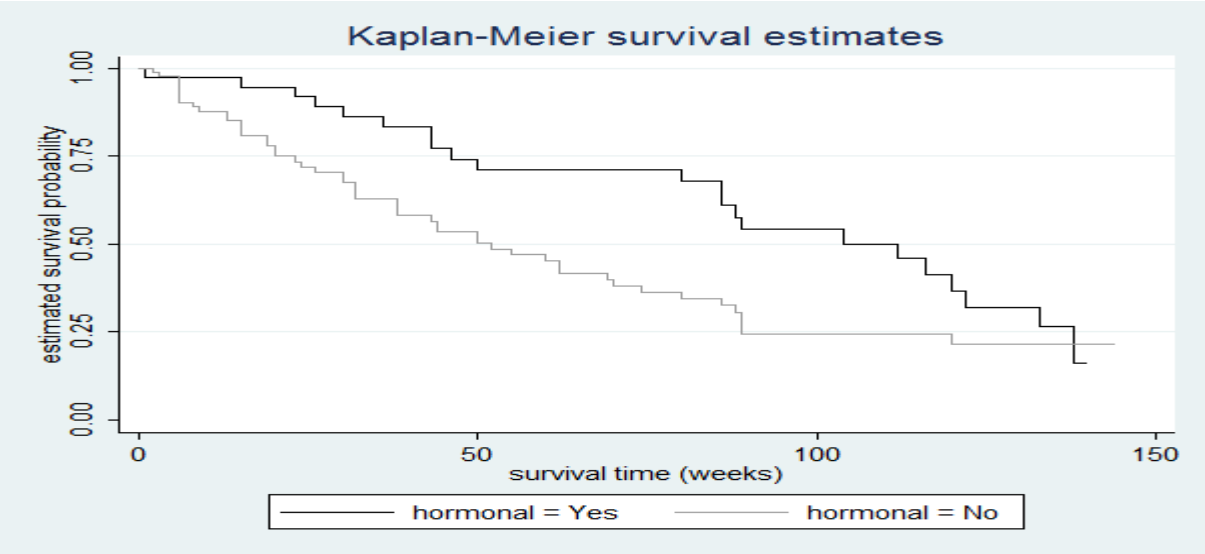
Log-rank test for equality of survival functions .4-7-1-8 :for Hormonal

Table (4-24) Log-rank test for equality of estimated survival functions for hormonal:

P-value	Chi-square test	Events expected	Events observed	Hormonal
0.0323	4.58	31.82	23	(Given (yes
		41.18	50	Not given ((no
		73.00	73	Total

Source: prepared by the researcher by using STATA, 2014

Figure (4-20): survival curves of patients hormonal



Source: prepared by the researcher by using STATA, 2014

We note from the table (4-24) that the value of chi-square test was (4.58), and the significant value to it (P-value = 0.0323 <0.05), there is significant difference between the estimated survival functions for hormonal. We note from figure (4-20) there is difference between the curves of Kaplan - Meier survival functions for hormonal. Thus, the probability of survival vary according to hormonal.

:**Estimate cox model for proportional hazards .4-7-2**

will be modeling univariate cox model for all significance variables in log-rank test, If will be significant model, it will be .included in multivariate cox model

:**Estimate cox model for stage .4-7-2-1**

:Table (4-25): Coefficient

[Conf. Interval 95%]		P>z	Z	.Std. Err	Coefficient	Factor
113156.-	.-5397593	0.003	3.00-	1088294.	3264576.-	Stage

Source: prepared by the researcher by using STATA, 2014

:Table (4-26): Hazard Ratio

[Conf. Interval 95%]		.Std. Err	Hazard Ratio	Factor
0.8971483	0.5856768	0.0788595	0.7248717	Stage

Source: prepared by the researcher by using STATA, 2014

:Table (4-27): Chi-square test

P-value	Chi-square test	Factor
0.0037	8.44	Stage

Source: prepared by the researcher by using STATA, 2014

We note from the table (4-25), that the significant value of Wald test p-value=0.003 <0.05), So estimated coefficient for the) variable is significant. That means the difference in the log hazard between stages is -0.326. , and it does not at least 0.586 and not more than 0.897 with 95% confidence interval From table (4-26) the hazard ratio is 0.725, means that at any time during the study, the per-week rate of death among .stage is 0.725 that of stage 0

We note from the table (4-27), that the univariate Cox model .(for stage is significant (p-value=0.0037 < 0.05

:**Estimate cox model for radiotherapy .4-7-2-2**

:Table (4-28): Coefficient

[Conf. Interval 95%]		P>z	Z	.Std. Err	Coefficient	Factor
1.081631	0541606.	0.030	2.17	2621146.	5678959.	Radiotherapy

Source: prepared by the researcher by using STATA, 2014

:Table (4-29): Hazard Ratio

[Conf. Interval 95%]		.Std. Err	Hazard Ratio	Factor
2.89713	1.037819	0.4541167	1.733983	Radiotherapy

Source: prepared by the researcher by using STATA, 2014

:Table (4-30): Chi-square test

P-value	Chi-square test	Factor
0.0435	4.08	Radiotherapy

Source: prepared by the researcher by using STATA, 2014

We note from the table (4-28), that the significant value of Wald test p-value=0.030<0.05), So estimated coefficient for the) variable is significant. That means the difference in the log hazard between radiotherapy is 0.568 , and it does not at least 0.054 and not more than 1.082 with 95% confidence interval. From table (4-29) A hazard ratio is 1.734, means that at any time during the study, the per-week rate of death among not .given radiotherapy is 1.734that of given radiotherapy
We note from the table (4-30), that the univariate Cox model for Radiotherapy .(is significant (p-value=0.0435 < 0.05

:**Estimate cox model for chemotherapy .4-7-2-3**

:Table (4-31): Coefficient

[Conf. Interval 95%]		P>z	Z	.Std. Err	Coefficient	Factor
1.397247	4342965.	0.000	3.73	2456551.	9157716.	Chemotherapy

Source: prepared by the researcher by using STATA, 2014

:Table (4-32): Hazard Ratio

[Conf. Interval 95%]		.Std. Err	Hazard Ratio	Factor
3.961492	1.519729	0.5997083	2.453649	Chemotherapy

Source: prepared by the researcher by using STATA, 2014

:Table (4-33): Chi-square test

P-value	Chi-square test	Factor
0.0002	13.94	Chemotherapy

Source: prepared by the researcher by using STATA, 2014

We note from the table (4-31), that the significant value of Wald test p-value=0.000<0.05), So estimated coefficient for the) variable is significant. That means the difference in the log hazard between chemotherapy is 0.568 , and it does not at least .0.054 and not more than 1.082 with 95% confidence interval From table (4-32) A hazard ratio is 2.454, means that at any time during the study, the per-week rate of death among not .given chemotherapy is 2.454 that of given chemotherapy We note from the table (4-33), that the univariate Cox model .(for chemotherapy is significant (p-value=0.0002 < 0.05

:**Estimate cox model for surgical .4-7-2-4**

:Table (4-34): Coefficient

[Conf. Interval 95%]		P>z	Z	.Std. Err	Coefficient	Factor
1.888217	0659883.	0.036	2.10	4648627.	9771025.	Surgical

Source: prepared by the researcher by using STATA, 2014

:Table (4-35): Hazard Ratio

[Conf. Interval 95%]		.Std. Err	Hazard Ratio	Factor
6.525027	1.05521	1.219575	2.623981	Surgical

Source: prepared by the researcher by using STATA, 2014

:Table (4-36): Chi-square test

P-value	Chi-square test	Factor
0.0170	5.69	Surgical

Source: prepared by the researcher by using STATA, 2014

We note from the table (4-34), that the significant value of Wald test p-value=0.0336 < 0.05), So estimated coefficient for the) variable is significant
From table (4-35) A hazard ratio is 2.624, means that at any time during the study, the per-week rate of death among did .not undergo surgery is 2.624 that of underwent surgery
We note from the table (4-36), that the univariate Cox model .(for surgical is significant (p-value=0.017 < 0.05

:**Estimate cox model for hormonal .4-7-2-5**

:Table (4-37): Coefficient

[Conf. Interval 95%]		P>z	Z	.Std. Err	Coefficient	Factor
1.037432	0379443.	0.035	2.11	254976.	537688.	Hormonal

Source: prepared by the researcher by using STATA, 2014

:Table (4-38): Hazard Ratio

[Conf. Interval 95%]		.Std. Err	Hazard Ratio	Factor
2.806177	1.032774	0.4341075	1.702394	Hormonal

Source: prepared by the researcher by using STATA, 2014

:Table (4-39): Chi-square test

P-value	Chi-square test	Factor
0.0321	4.59	Hormonal

Source: prepared by the researcher by using STATA, 2014

We note from the table (4-37), that the significant value of Wald test p-value=0.035 < 0.05), So estimated coefficient for the) variable is significant
From table (4-38) A hazard ratio is 1.702, means that at any time during the study, the per-week rate of death among not given hormone therapy is 1.702 that of given hormone .therapy
We note from the table (4-39), that the univariate Cox model .(for hormone therapy is significant (p-value=0.0321 < 0.05

:**Estimate multivariate cox model .4-8**

:Table (4-40): Coefficients

[Conf. Interval 95%]		P>z	Z	.Std. Err	Coefficient	Factor
0639861.-	494122.-	0.011	2.54-	1097306.	2790541.-	Stage
9684132.	0831222.-	0.099	1.65	2682538.	4426455.	Radiotherapy
1.32268	3281814.	0.001	3.25	2537034.	8254309.	Chemotherapy
1.949584	091165.	0.031	2.15	4740952.	1.020375	Surgical
1.396005	3447637.	0.001	3.25	2681787.	8703843.	Hormonal

Source: prepared by the researcher by using STATA, 2014

:Table (4-41): Hazard Ratios

[Conf. Interval 95%]		.Std. Err	Hazard Ratio	Factor
9380181.	6101063.	0830111.	756499.	Stage
2.633762	9202387.	4176229.	1.55682	Radiotherapy
3.753469	1.388441	5791704.	2.282864	Chemotherapy
7.025765	1.09545	1.315251	2.774234	Surgical
4.039032	1.411656	6403648.	2.387828	Hormonal

Source: prepared by the researcher by using STATA, 2014

:Table (4-42): Chi-square test

P-value	Chi-square test	Model
0.0000	37.02	

Source: prepared by the researcher by using STATA, 2014

From table (4-42) the multivariate Cox model is significant (p-value=0.000 < 0.05).

We note from the table (4-40), that there is significant value of Wald test for: (1) Stage (p-value=0.011<0.05), (2) Chemotherapy (p-value=0.001 <0.05), (3) Surgical (p-value=0.031<0.05), (4) Hormonal (p-value=0.001 <0.05). So estimated coefficients for these variable is significant. It will be included in the multivariate model as follows:

From table (4-41) A hazard ratio for: (1) Stage is 0.756, means that at any time during the study, the per-week rate of death among stage is 0.756 that of stage 0, (2) Chemotherapy is 2.283, means that at any time during the study, the per-week

rate of death among not given chemotherapy is 2.283 that of given chemotherapy, (3) Surgical is 2.774, means that at any time during the study, the per-week rate of death among did not undergo surgery is 2.774 that of underwent surgery, (4) Hormonal is 2.389, means that at any time during the study, the per-week rate of death among not given hormone therapy is 2.389 that of given hormone therapy.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

.5.1 Preface:

This chapter contains the results that have been reached through the practical side of the research, in addition to the proposed recommendations.

.5.2 Conclusions:

1- There are significant differences on hazard ratio among patients who given the chemotherapy and who did not given chemotherapy (P-value =0.000 < 0.05).

2- There is a significant difference to the hazard ratio in terms of the stage of the disease (P-value =0.003 < 0.05).

3- There are significant differences between patients who given radiotherapy and those who did not given radiotherapy in terms of hazard ratio (P-value =0.030 < 0.05).

4- There are significant differences between patients who given hormonal therapy and those who did not given

hormonal therapy in terms of hazard ratio (P-value =0.035 .(< 0.05

There are significant differences on hazard ratio among -5 patients who have undergone surgery, and those who did .(not undergo surgery (P-value =0.036 < 0.05

.5.3 **:Recommendations**

:The study recommended the following

The possibility of using the Cox model for the hazard ratio - 1 of multivariate in the calculation of the value of hazard at any .given time

The use of parametric survival regression models in the - 2 .analysis of the study data

References

Arkoob K., AL-Nsour M., AL-Nemry O., AL-Hajawi B. (2010), -1 Epidemiology of breast cancer in women in Jordan: patient characteristics and survival analysis, Eastern .Mediterranean Health Journal Vol.16 No.10

2- Collett D., 2003. Modeling Survival Data in Medical Research. Chapman and Hall London.

3- Cox, D.R, and Snell,E.J.A general definition of residuals with discussion. Journal of the Royal Statistical Society series B30 (1968).

4- Cox, D.R, Regression Models and life-tables. Journal of the Royal Statistical Society series B34 (1972).

5- Dakhil Nazera Khalil, AL-Decemberali Yahya Mahdi, AL-Abidy Muna Abbas (2012), Analysis of Breast Cancer Data

using Kaplan–Meier Survival Analysis, Journal of Kufa for Mathematics and Computer Vol.1 , No.6, P.7-14.
6- David W. Hosmer, Stanley Lemeshow (2007). Applied Survival Analysis: Regression Modeling of Time-to-Event Data. 2nd edition Wiley, New York.
7- Deshpande Anjali D., Jeffe Donna B., Gnerlich Jennifer, Iqbal Ayesha Z., Thummalakunta Abhishek and Margenthaler Julie A. (2009), Racial Disparities in Breast Cancer Survival: An Analysis by Age and Stage, Journal of Surgical Research Vol. 153, No. 1, P. 105–113.
8- Klein, John P. and Moeschberger.L.2003 Survival Analysis Techniques for censored and truncated data. 2rd edition New York.
9- Scientific article prepared by D. Mahmoud Shaheen and others.
10- Therneau, T. and Grambsch, P.2000. Modeling Survival Data; extending the Cox model.
11- Wei Lj (1992) The accelerated failure time models useful alternative to the Cox Regression model in survival analysis.
12- Ziaei Jamal Eivazi, Sanaat Zohreh, Asvadi Iraj, Dastgiri Saeed, Pourzand Ali and Vaez Jalil (2013), Survival Analysis of Breast Cancer Patients in Northwest Iran, Asian Pacific Journal of Cancer Prevention, Vol. 14, P. 39-42.