STATISTICAL ASSESSMENT OF CERVICAL SQUAMOUS CELL CARCINOMA ATTRIBUTABLE RISK TO HIV-INFECTED WOMEN IN ZAMBIA

By

GILLIEN CHISALA

A thesis submitted to the University of Zambia, in partial fulfilment of the requirement for the Master of Science Degree in Statistics in the Department of Mathematics and Statistics, School of Natural Sciences.

THE UNIVERSITY OF ZAMBIA

2016

Supervisor: Dr. A. Banda

AUTHOR'S DECLARATION

I, Chisala Gillien, do declare that this dissertation represents my own work and that it has neither in part nor in whole, been presented as substance for the award of any degree at this or any other institution of learning or research. Where other people's work has been used, acknowledgement has been made.

Signature of Author:…….....

Date: …….....

Signature of the Supervisor: …….....

Date: …….....

CERTIFICATE OF APPROVAL

The University of Zambia approves the dissertation of Chisala Gillien as fulfilling part of the requirements for the award of the degree of Master of Science in Statistics.

Name:……

Signature:……

Date:……

Name:……

Signature:……

Date:……

Name:……

Signature:……

Date:……

Name:……

Signature:……

Date:……

Copyright

DEDICATION

This humble piece of my research is dedicated to my family.

My mother and father, Mr and Mrs Chisala:

You have seen me this far; words alone cannot express how grateful I am to you. Thank you for your ever encouraging words and support. You have been the pillar of my strength, all that I am and hope to be I owe to you.

My brothers and sisters:

Thank you for understanding, encouragement and moral support during the time of conducting this study.

And most of all, to my Creator and God, he has been with me all these years and it is by His Power and Grace that I have come this far. Without Him, all is in vain.

ACKNOWLEDGEMENTS

This thesis would be incomplete without thanking the following people who have contributed much to the success of this work and stay at the University. First and foremost I would like to thank the Lord Almighty for his assuring love and care. **"Small and weak as you are, don't be afraid; I will help you. I, the Holy God of Israel, am the one who saves you"** Isaiah 41:14

My profound gratitude and thanks go to Dr. A. Banda for the expertise and intellectual guidance he displayed during the proof reading of the manuscript of this report. Thank you for sacrificing so much of your time in giving me direction, may God Richly reward you for your tireless effort. Special thanks go to Mr. Tembo H, for his patience and intellectual guidance during the data processing exercise and eventual production of this report.

Once more I wish to thank Modester who have made this research a success.

My thanks are also due to the Ministry of Higher Education, Department of Science and Technology for financial assistance.

ABSTRACT

Cervical cancer is the second most common cancer among women worldwide with estimated 529,409 new cases and 274,883 deaths in 2008, as reported by the International Agency for Research on Cancer. In Zambia, cervical cancer is estimated to be the leading cause of cancer and cancer mortality in women of all ages. Research does seem to indicate that being infected with the human immunodeficiency virus (HIV) increases susceptibility to the most common cancer cell, the squamous cell carcinoma. The magnitude of the risk, however, is unknown for the case of Zambia. To determine the attributable risk of this cell on HIV positive subjects the study utilized a matched case-control design on data from medical records at National Cancer Diseases Hospital (NCDH) in Lusaka for period 2007 to 2010. Medical records of women aged 18 to 59 years old and whose information on age, marital status, HIV status and cancer diagnosis were available constituted the target population. Analysis from a sample of 348 such subjects indicates a significant association between cervical squamous cell carcinoma and HIV, a Chi-square test of one degree yielded a value of 5.63. The risk of having cervical squamous cell carcinoma attributed to HIV was estimated to be 2.1542 times higher for women exposed HIV as compared to those unexposed. The findings showed that HIV-infected women are at higher risk of cervical squamous cell carcinoma than HIV-infected women and there is strong association between CSCC and HIV among women in Zambia.

LIST OF TABLES

OPERATIONAL DEFINITIONS OF KEY CONCEPTS AND TERMS

Association – it is the extent to which the occurrence of two or more characteristics is linked either through a causal or non-causal relationship.

Case – it is a subject with the disease of interest.

Case-Control Study – it is an observational study in which subjects are sampled based on the presence (cases) or absence (controls) of the disease of interest.

Control – it is the subject without the disease of interest.

Cumulative Incidence – is the risk of developing a particular disease within a specified period of time.

Excess Risk – the extra risk of a particular disease occurring among persons exposed to a risk factor of interest.

Exposure – contact with or possession of a characteristic that is suspected to influence the risk of developing a particular disease.

Incidence Rate – it is the rapidity with which new cases of a particular disease arise within a given population.

Matching – a procedure for sampling comparison subjects based upon whether key attributes are similar to those of subjects in the index group.

Odds – the probability that a particular event will occur divided by the probability that the event will not occur.

Odds Ratio – the odds of a particular exposure among the subjects with a specific disease divided by the corresponding odds of exposure among subjects without the disease of interest.

Prevalence – the proportion of subjects in a given population who have a particular disease at a point or interval of time.

Rate – the rapidity with which health events such as new diagnoses or deaths occur.

Risk – the probability that an event will occur within a specific period of time.

Risk Factor – an attribute or agent that is suspected to be related to the occurrence of a particular disease.

Sample – a subject of a target population that is chosen for investigation.

Type I Error – rejection of the null hypothesis when it is actually correct.

Type II Error – failure to reject the null hypothesis when it is actually incorrect.

LIST OF ABBREVIATIONS

AIDS - Acquired Immunodeficiency Syndrome.

ART- Antiretroviral Therapy.

CAP - College of American Pathologist.

CCC - Cervical Cancer Clinics.

CDC - Centers for Disease Control.

CIN - Cervical Intraepithelial Neoplasia.

HIV- Human Immunodeficiency Virus.

HPV - Human Papilloma virus.

IARC - International Agency for Research on Cancer.

ICC - Invasive Cervical Cancer.

NCDH - National Cancer Diseases Hospital.

SIL - Squamous Intraepithelial Lesions.

SOGC - Society of Obstetricians and Gynaecology of Canada.

STI's - Sexually Transmitted Infections.

VIA - Visual Inspection with Acetic Acid.

WHO - World Health Organisation.

CHAPTER 1: INTRODUCTION

1.0 Background

Cervical cancer is cancer that starts in the cervix. The cervix is the lower part of the uterus (womb); the cervix is sometimes called the uterine cervix. It has two main types of cells and these are the squamous cells which are found on the exocervix and the glandular cells which are found on the endocervix. The place where the exocervix and endocervix meet is called the transformation zone, and this is where most cervical cancers start, in the cells found in the lining of the cervix. Before the cancer cells develop, there is a gradual change from normal healthy cells to pre-cancerous cells which later develop into cancer cells. When these cells change, cancer of the cervix occurs and affects deeper cell layers or spread to other organs near the cervix.[1]The majority of women who have pre-cancer cells need no treatment, as these will go away naturally. However, since there is the possibility that pre-cancer cells could develop into cancer, they should seek treatment.

1.1 Types of cervical cancer

There are many types of cervical cancer; the most common ones are squamous cell carcinoma and adenocarcinoma. They are distinguished based on their appearance under a microscope. According to Khan et al., 2005, both squamous cell carcinoma and adenocarcinoma begin in the cells that line hollow organs; squamous cells have a thin flat appearance while adenocarcinomas involve cells with secretory functions.

The squamous cell carcinoma develops from the squamous cells (flat cells) that cover the outer surface of the cervix (the ectocervix) at the top of the vagina.

Adenocarcinomas cancers arise from glandular or columnar cells, they start on the endocervix, primarily with the cells that produce mucus. The less common type of cervical cancers are known as adenosquamous carcinomas or mixed carcinomas and have features of both squamous cell carcinomas and adenocarcinomas.[2]

[1] F. James, 2008.
[2]American Cancer Society, 2010.

1.2 Cervical cancer situation analysis

In Zambia, cervical cancer is estimated to be the leading type of cancer and cancer mortality in women of all ages. In 2002, the incidence estimates indicated that 1,650 women were diagnosed with cancer of the cervix and that 1,340 (81.2%) died from the disease. In 2008, the incidence estimates were that every year 1,839 women are diagnosed with cervical cancer and 1,276 (69.4%) die from the disease. The reduction could be as a result of sampling variation or under estimation of incidence. The increase in death rates could be as a result of not going early for cervical cancer screening.

1.3 Statement of the problem

During the past decade, HIV infected women in low socio- economic settings like Zambia had very limited access to antiretroviral therapy, consequently, had a very short survival period after being diagnosed with HIV. However, increasing numbers of women are now on antiretroviral therapy treatment programs that have the potential to improve their lifespan for a considerable period.[3]

Despite the fact that now, HIV positive women are accessing antiretroviral therapy treatment, they still have short survival. This is because HIV acts as a co-factor for the persistent oncogenic human papilloma virus (HPV) infection, which plays the causal role in the development of cervical neoplasm which is a sufficient criterion for AIDS, even in the absence of an opportunistic infection. In the presence of squamous carcinoma lesions, the problem is worsened though it is considered to be a preventable condition. This is because it is associated with a long pre-invasive stage, making it amenable to screening and treatment as long as it is detected early and managed.[4]

However, the risk squamous cell carcinoma poses to HIV-infected women compared to HIV-uninfected women has not been ascertained in the case of Zambia. Hence, the need to determine the excess risk of squamous cell carcinoma of the cervix in HIV-infected women compared to HIV-uninfected women.

[3]G.P. Parham. Prevalence and predictors of squamous intraepithelial lesions of the cervix in HIV-infected women in Lusaka, Zambia, (103) 2006 pp. 1017-1022.
[4]D.M. Parkin, P. Pisani, J. Ferlay. Estimates of the world of 25 major cancers, Int J Cancer 1990 pp. 827-841.

1.4 Aim of the study

The aim of this study was to determine whether squamous cell carcinoma of the cervix does pose higher risk among HIV-infected women as compared to HIV-uninfected women and to determine the magnitude of that risk.

1.5 Specific objectives

The research objectives of this study are to:

(1) Determine whether or not squamous cell carcinoma of the cervix poses higher risk among HIV-infected women.

(2) Determine the excess risk of exposure to cervical squamous cell carcinoma among HIV positive women.

1.6 Significance of the study

The significance of the study is that the findings may lead to screening for cervical squamous cell in HIV-infected women which can improve the survival in HIV-infected women once diagnosed by treating squamous cell carcinoma. It can also lead to the knowledge of how much of the risk is present in HIV-infected women in Zambia. Finally this may lead to the need of designing appropriate treatment programs in Zambia.

1.7 Hypothesis

The incidence rate of cervical cancer is higher in HIV-infected women than HIV-uninfected women.

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

This chapter presents some facts about cervical squamous cell carcinoma, its burden worldwide and in Zambia, its risk factors, its association with HIV, its prevention, role of screening, attributable risk or excess risk and gap analysis arising from previous studies.

2.1 Burden of cervical squamous cell carcinoma

Atashili (2009) stated that cervical cancer is generally used in reference to squamous cell carcinoma of the uterine cervix. It is the second most common cancer among women worldwide with annual incidence and death rate of 471,000 and 233,000 respectively.[5] Worldwide, the incidence of the disease is declining but 80% of all cases are found in developing countries with the highest rates being in Latin America and Southeast Asia. These high rates persist despite the existence of a proven cost-effective screening technique, the pap smear.

In Zambia, cervical cancer is estimated to be the first leading type of cancer and cancer mortality in women of all-ages (1,839 women are diagnosed with cervical cancer annually and 1,276 die from the disease).[6]This is because there are fewer screening programs for early detection of pre-cancerous lesions within the country.

2.2 Risk factors for cervical squamous cell carcinoma

A risk factor is any agent that increases your chance of getting a disease such as cancer. Different cancers have different risk factors. Being exposed to one or more risk factors does not mean that you will get the disease. However, several risk factors increase the chance of developing the disease.

Persistent infection with oncogenic HPV is associated with the development of cervical squamous cell carcinoma.[7] There are varying magnitude of associations, these include factors

[5]D.M. Parkin. Global cancer statistics in the year 2000.Lancet Oncology (2) 2001 pp. 533-543.
[6]GLOBOCAN 2008.
[7]CAP, 2011.

such as family history on cervical cancer, multiple pregnancies, viral infections, early onset of sexual activity, multiple sex partners, smoking and long term use of oral contraceptives.[8]

2.2.1 Human Papilloma Virus (HPV)

HPV is the most important risk factor for cervical squamous cell carcinoma.[9]It is passed from one person to another during skin-to-skin contact with an area of the body infected with HPV .[10]

2.2.2 Multiple sex partners or early onset of sexual activity

Cervical squamous cell carcinoma behaves like a sexually transmitted disease. Several indicators, the most convincing and consistent being multiple sexual partners and young age at first intercourse.[11] So age is crucial in this study because during puberty, cervical tissue undergoes a variety of changes that may make the area more vulnerable to damage.

2.2.3 Family history on cervical cancer

Cervical cancer may run in some families. If the mother or sister had cervical cancer, then one's chances of developing the disease are 2 to 3 times higher than if no one in the family had it.

2.2.4 Multiple pregnancies

Women who have had 3 or more full-term pregnancies have an increased risk of developing cervical squamous cell carcinoma.[12]This means that cervical squamous cell carcinoma behaves like a sexually transmitted disease. This indicates that marital status has an impact on cervical squamous cell carcinoma.

2.2.5 Smoking

Research indicates that women who smoke are about twice as likely as non-smokers to get cervical squamous cell carcinoma.

2.2.6 Long term use of oral contraceptive

Research also indicates that the use of oral contraceptive for 5 years or more increases the chances of having cervical cancer.

---

[8]Green et al. "Risk factors for adenocarcinoma and squamous cell carcinoma of the cervix in women aged 20-44 years: the UK National Case-Control Study of Cervical Cancer", *British Journal of Cancer* 89 (2003), pp.2078-2086.
[9] CAP, 2011
[10]American College of Obstetricians and Gynaecologists."Human papilloma virus infection," 1951.
[11]American Cancer Society, 2007.
[12]P. Beral et al."Cervical Cancer and Hormonal Contraceptives," 2007.

2.3 Cervical squamous cell carcinoma and squamous carcinoma lesions in HIV-infected women

Cervical cancer is the second most common malignancy and accounts for the greatest number of deaths from cancer in women worldwide.[13] Human immunodeficiency virus (HIV) infection also represents a tremendous health burden worldwide. The HIV/AIDS pandemic may have contributed to the high incidence of cervical cancer in Zambia, where the number of cases is the second highest in sub-Saharan Africa, and HIV prevalence is one of the highest in the world.

In HIV-infected women, there is an increased risk of HPV infection and squamous intraepithelial lesions (SIL), the precursor of cervical cancer.[14]According to research findings, it has been suggested that HIV-induced immunodeficiency predisposes to cervical intraepithelial neoplasia (CIN) or cervical carcinoma by facilitating the expression of a causal agent.[15]

2.4 Prevention of cervical squamous cell carcinoma

Prevention of cervical squamous cell carcinoma is the way of stopping or preventing the disease from occuring. Since HPV is the main cause of cervical squamous intraepithelial lesions, the primary prevention rests in the prevention of HPV infection. This can be done by avoiding contact with HPV.[16] Although avoiding exposure to HPV could help prevent this disease, more specific HPV targeted intervention is vaccination.

Secondary prevention is aimed at the early detection of pre-cancerous lesions and their treatment. Various guidelines exist depending on the target population and the issuing agency. Women at any age should not be screened annually by any screening method; rather, recommended screening intervals for women are based on age and clinical history.[17]

Cervical screening is a way of checking for squamous intraepithelial lesions (SILs) of the uterine cervix. Organised screening does not only prevent morbidity and costs of treating advanced

[13]D. M. Parkin et al. Global Cancer Statistics, 2011.

[14]C. S. Rabkin et al. "Cervical incidence trends in women at high risk of human immunodeficiency virus (HIV) infection," International Journal of Cancer, 55 (1993), pp. 208-212, X. W. Sun et al. "Human papilloma virus infection in women infected with human immunodeficiency virus," New England Journal of Medical, 337 (1997), pp. 1343-1349.

[15] A. R. Feingold et al. "Cervical cytology abnormalities and papilloma virus virus in women infected with human immunodeficiency virus, Journal of Acquired Immune Deficiency Syndromes, 9(1990)

[16] J. Atashili. "Cervical Precancerous Lesions in HIV-infected women in Cameroon: Prevalence, Predictors and Potential impact of Screening," 2009.

[17] J. Atashili, 2009.

disease, it also reduces the deaths due to cervical cancer by up to 80%.[18] Thus, the fundamental goal of screening is to prevent morbidity and mortality from cervical cancer.[19]

There are, at least, two critical components of a cervical screening program, coverage and follow up activities. Several reports have stated that ensuring an adequate follow up of women with Pap smears abnormalities is very important to decrease the incidence of the disease and deaths due to it.[20]

Screening coverage is commonly used as the only indicator to evaluate the success of cervical screening programs. However, screening is just a part of them, therefore screening programs must be evaluated for the full range of services needed to prevent cervical squamous cell carcinoma and reduce its mortality.

2.5 Attributable risk/excess risk

Measures of impact are used to assess the contribution of one or several exposures to the occurrence of incident cases at the population level. The most commonly used measure of impact is the attributable risk.[21] Attributable risk (AR) is a measure of excess risk that is attributed to a specific exposure. It is an important tool for determining how much of an outcome may be attributable to a particular risk factor (i.e. an estimate of the excess risk) in a population exposed to that factor.[22] AR is valuable in public health to weigh the impact of exposure on the burden of disease occurrence and assess potential impact of prevention programs aimed at reducing or eliminating exposure from the population. Attributable risk is often thought of as the fraction of disease that could be eliminated if exposure could be totally removed from the population. AR depends both on the strength of the association between exposure and disease and the prevalence of exposure in the population. This means that attributable risk increases both with the strength of the association between exposure and disease measured by relative risk, and with the prevalence of exposure in the population.

In this study, the AR will help to determine how much of the cervical squamous cell carcinoma risk experienced by HIV-infected women can be attributed to HIV. In practice, it implies how

[18] National Cancer Institute, 2010.
[19] L. Denny. Prevention of cervical cancer. Reproductive Health Matters, 16 (2008), pp. 18-31.
[20] L. P. Engelstad et al. "The effectiveness of a community outreach intervention to improve follow-up among underserved women at higher risk for cervical cancer", 41 (3-4) 2005, pp. 741-748.
[21] Benichou, J. (2007). Measuring Attributable Risk. HRSA
[22] Rosenberg, D. and Handler, A. (1998). Measures of Association and Hypotheses

much of the risk (incidence) of disease we can hope to prevent if we reduce or eliminate exposure to HIV.

2.6 Gap analysis arising from previous studies

There is a lot of literature focusing on prevalence and other related statistics on cervical squamous intraepithelial lesions, the following studies have some similarities with this research:
Adjorlolo-Johnson et al. (2010) did a study on assessing the relationship between HIV-infection and cervical cancer in Cote d'lvoire. The aim was to assess the relationship between HIV-infection and invasive cervical cancer taking into account HPV infection and other potential risk factors for cervical cancer. The study revealed that HIV infection was associated with cervical cancer in women with HPV infection. Among women aged less or equal to 40 years the risk factors for cervical cancer were high risk HPV infection, parity greater than two and HIV-infection. However, the association of cervical cancer with HIV-infection in women aged over 40 years, the high risk HPV infection and parity greater than two was not statistically significant.

In Tanzania, Kapiga et al. (2009) researched on prevalence of lesions whose aim was to determine the prevalence and risk factors for squamous intraepithelial lesions in HIV-infected women. The study revealed that the prevalence of SIL was 2.9%; Low Squamous Intraepithelial Lesions (LSIL) was 1.6% while High Squamous Intraepithelial Lesions (HSIL) stood at 1.3% and the number of lifetime sexual partners and parity were marginally and non-significantly associated with SIL.

In Zambia, a similar study was conducted in Lusaka by Parham et al. (2006) whose aim was to evaluate the prevalence and predictors of squamous intraepithelial lesions in HIV-infected women aged between 23 and 49. The prevalence was found to be 56% with an Odds Ratio (OR) of 12.4 in HIV-infected women. While another study done by Sahasrabuddhe et al. (2007) on prevalence of HPV and lesions reported a lower prevalence with an OR of 8.0. The difference in the OR could have been due to a number of factors, which could include, differences in target populations used, study period and age groups selected in the two studies.

However, the above studies focused on the association between cervical cancer and HIV-infected women, on the prevelence of cervical squamous intraepithial lesions among HIV-infected women whose prevelence and OR were reported. However, one of the gaps that exist, to mention

but one, is the ascerntainment of the burden of risk of squamous cell carcinoma of the cervix among HIV positive women. Therefore, in filling the gap, the study aimed at determining whether or not squamous cell carcinoma of the cervix poses higher risk among HIV-infected women. The excess risk of exposure to cervical squamous cell carcinoma among HIV positive women was also determined.

CHAPTER 3: METHODOLOGY

3.0 Introduction

The chapter describes the methodology that was used in the study which includes among other things, study design, data collection, and description of variables, sample size determination and analysis.

3.1 Study design

In this research, we were looking at the association between cervical squamous cell carcinoma and exposure to HIV. The subjects in this study were women with cervical cancer of different cell types. The association between cervical squamous cell carcinoma and exposure to HIV were determined using a case-control design. In this study, any subject with squamous cell carcinoma was treated as a case and any other subject with different cervical cancer cell type was treated as a control. The risk factor of interest is exposure to HIV which had two levels; HIV positive (HIV+) or HIV negative (HIV-). The response of interest is the status of squamous cell carcinoma which also had two levels; a case or a control. After data collection, the information on the response variable and risk factor was represented in a table such as the one shown below.

Table 1: A typical cross tabulation of a response and a risk factor.

		Response outcome		
		Cases	Controls	Total
Exposure	HIV+	a	b	m_1
	HIV-	c	d	m_2
	Total	n_1	n_2	n

Initially before analysis, all we knew was that we had n_1 cases and n_2 controls. HIV status was determined for each case as well as for each control.

3.1.1 Matching

Literature does suggest that age and multiple pregnancies have a bearing on cervical cancer, consequently, in this study, the subjects were matched on age and marital status. The process of matching is discussed below.

To avoid sparse data when age and marital status were matched, age was converted into a categorical variable with four levels; 18-29, 30-39, 40-49 and 50-59 years of age.

The data used covered a period from 2007 to 2010. The rates of cervical cancer may vary from year to year and to account for this variation, matching was done for each year.

Below is a dummy table showing a typical cross tabulation of year, age category and marital status.

Table 2: A dummy table of a cross tabulation on age category and marital status.

YEAR	AGE CATEGORY	MARTIAL STATUS			TOTAL
		Single	Married	Divorced/widowed	
20XX	18 – 29	a	b	c	n_1
	30 – 39	d	e	f	n_2
	40 – 49	g	h	i	n_3
	50 – 59	j	k	l	n_4

In each cell, an equal numbers of cases and controls were randomly selected from the data collected. This was done for each year separately, as long as the cells have sufficient numbers. From the selected cases and controls, HIV status was then determined.

3.1.2 Measure of risk

Common measures of association between disease and exposure in a Case-Control study include Incidence Density Ratio (IDR), Cumulative Incidence Ratio (CIR), Exposure Odds Ratio (EOR), but the latter is the most commonly used. In this study, the EOR was used to determine the association between squamous cell carcinoma and exposure to HIV.

In Table 1, let D+ indicate a subject who is a case and D- indicate a subject who is a control. Further, let E+ indicate a subject who is HIV+ and E- indicate a subject who is HIV-. Then;

(i) The odds of exposure among cases is defined as: $odds_1 = \frac{Pr(E+/D+)}{Pr(E-/D+)}$,

When data are available and presented as in Table 1, the estimate of $odds_1 = \frac{a/n_1}{c/n_1} = \frac{a}{c}$

(ii) The odds of exposure among controls is defined as: $odds_2 = \frac{Pr(E+/D-)}{Pr(E-/D-)}$.

The estimate is given by $odds_2 = \frac{b/n_2}{d/n_2} = \frac{b}{d}$

(iii) The estimate of the exposure odds ratio, EOR, is then defined as:

Odds = $\psi = \frac{Odds_1}{Odds_2} = \frac{ad}{bc}$,

(iv) Let In Ψ be the log of EOR, then the estimate of the standard error for In Ψ is given by:

$$SE(\text{In } \Psi) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

3.2 Data collection procedure

Below is a typical unedited patients records showing dummy registration number (RegN), sex of a patient (sex), age, marital status (Mari), site, HIV status (HIV) and diagnosis.

Patients Records

RegN	Sex	Age	Mari	Site	HIV	Diagnosis
2008000X	2	23		539	2	Hemangiosarcoma of Cervix
2008000X	2	9	1	419	2	Ewing Sarcoma
200800XX	2	29	2	779	2	Kaposi Sarcoma of Lymph node
200800XX	1	55	2	619	2	Adenocarcinoma of Prostate
200800XX	1	67	2	779	1	Kaposi Sarcoma of Lymph node
200800XX	1	54	2	619	2	Adenocarcinoma of Prostate
200800XX	2	61	2	539	2	Non-keratinising Squamous Cell Carcinoma

To create data of a case-control study, all cases of cervical cancer were extracted from patient's records from 2007 to 2010 such as the one shown above. Each cancer type is given a specific Site number and cervical cancer is given Site number 539. All cervical cancer cases were extracted using Site number 539. Sex and Diagnosis as shown above were also used to verify cervical cancer records extracted. The study limits the age group from 18-59 years old. The patients' record came from National Cancer Diseases Hospital (NCDH).

3.3 Derived variables

3.3.1 Disease status

To create the response variable (disease status) with two levels (case and control), an SPSS function used the variable diagnosis and assign a subject to a case if the diagnosis is squamous cell carcinoma, or assign a subject to a control if diagnosis is any other cancer cell among subjects with cervical cancer.

3.3.2 Year

To create the variable year, an SPSS character function was used to extract the first four digits the registration number (RegN) from patients' records. The first four digits of the registration number represent a year. For example, in the patients' records shown above, the first four digits represent 2008.

3.4 Recorded variables

3.4.1 Sex

In the patients' records, sex is recorded and assigned 1 for males and 2 for females.

3.4.2 Age

For each patient, age is recorded in years, but for this study we created a categorical age variable with four levels and these were given earlier. This variable was created to guard against sparse data when matching was done.

3.4.3 Marital status

For each patient, marital status is recorded and in this study it had three levels and these were; single, married and divorced/widowed.

3.4.4 HIV status

HIV status is also recorded for each patient and an HIV positive subject is assigned 1 and an HIV negative subject is assigned 2. HIV was our exposure of interest and disease status was determined by the presence or absence of squamous cell carcinoma as described in section 3.3.1.

3.5 Sample size

It has been stated earlier that cases were subjects with cervical squamous cell carcinoma and controls were subjects without cervical squamous cell carcinoma but other types of cervical cancer cells. The exposure in this study is HIV and a subject who is HIV positive was assigned the code HIV+ and a subject who is HIV negative was assigned a code HIV-.

Suppose that;

(i) p_0 *is the proportion of women with squamous cell carcinoma who are HIV negative in the population.*

(ii) p_1 *is the proportion of women with squamous cell carcinoma who are HIV positive in the population.*

Then for a case control study, the formula for a sample size for each group (cases or controls) is given by;

$$\mathrm{n} = \frac{\left\{z_{1-\alpha/2}\sqrt{2pq} + z_{1-\beta}\sqrt{p_1(1-p_1)+p_0(1-p_0)}\right\}^2}{(p_1-p_0)^2}$$

Where

- n is the sample size of each group of cases and controls
- $1-\beta$ is the power of the test.
- α is the level of significance for the test or probability of type 1 error
- Z_r is the percentile value of the standard normal distribution, where: $r = 1-\frac{\alpha}{2}$ *or* $1-\beta$
- $p = \frac{p_1+p_0}{2}$, $q = 1-p$

p is the average of the proportion of women with and without HIV and have cervical squamous cell carcinoma.

Due to lack of information on the prevalence of cervical squamous cell carcinoma among HIV negative women the researcher set P_0 at 0.5. Using 56% as the prevalence of squamous intraepithelial lesions in HIV-infected women, reported by Parham el al (2006), the researcher rounded up the figure to 60%. Type I error was set at 5% and Type II error at 20% yielding the values of $Z_{1-\alpha/2} = 1.96$ and $Z_{1-\beta} = 0.842$, respectively. Inserting these values in the formula above yields a sample size of 388.

3.6 Statistical analysis

After selection of subjects from patients' record obtained from NCDH and matching on marital status and age, the data was tabulated in the format of Table 1, for each year. The study utilized the odds ratio to assess the association between disease and exposure.

The attributable risk (AR) was calculated to determine how much risk of cervical squamous cell carcinoma is due to HIV. The formula of calculating AR in a case-control study is given below.

3.6.1 Attributable risk

Suppose we have a case-control study, let AR represent the attributable risk. Then formula for AR is given by:

$$AR = \frac{p(IDR-1)}{p(IDR-1)+1}, \text{ and } p = \frac{b}{n_2}$$

Where Incidence Density Ratio (IDR) is estimated by the odds ratio and p is estimated from the target disease-free population, b and n_2 are as indicated in Table 1.

3.6.2 Pooled estimates

To account for year to year variation, a pooled estimates of odds ratio was obtained which was then used to assess the overall association between the disease and exposure using the Mantel-Haenszel statistic.

Table 3: A typical cross tabulation of disease status and exposure status.

Exposure status		Disease status		
		Cases	Controls	Total
Exposure Status	HIV+	a_i	b_i	m_{i1}
	HIV-	c_i	d_i	m_{i2}

Total	n_{i1}	n_{i2}	n_i

Assume Table 3 represents data for table 1 for year i, then:

1. Estimate of exposure odds ratio as discussed in section 3.1.2 is

   $\hat{\psi}_i = \frac{a_i d_i}{b_i c_i}$, i = 1, 2, 3 and 4.

2. Estimate of the standard error of the log of $\hat{\psi}_i$ is

   $$SE(ln\hat{\psi}_i) = \sqrt{\frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}}.$$

3. An estimate of a weight for year i in the pooled estimate of the log of of $\hat{\psi}_i$ is given by

   $$w_i = \frac{\left(\frac{1}{SE(ln\hat{\psi}_i)}\right)^2}{\sum_{i=1}^{4}\left(\frac{1}{SE(ln\hat{\psi}_i)}\right)^2}.$$

4. The pooled estimate of the log of $\psi_i$is given by

   $$ln\hat{\psi}_w = \hat{\theta} = \frac{\frac{ln\psi(2007)}{[SE(ln\psi(2007))]^2} + \frac{ln\psi(2008)}{[SE(ln\psi(2008))]^2} + \frac{ln\psi(2009)}{[SE(ln\psi(2009))]^2} + \frac{ln\psi(2010)}{[SE(ln\psi(2010))]^2}}{\frac{1}{[SE(ln\psi(2007))]^2} + \frac{1}{[SE(ln\psi(2008))]^2} + \frac{1}{[SE(ln\psi(2009))]^2} + \frac{1}{[SE(ln\psi(2010))]^2}}$$

3.6.3 Test for overall association

The null hypothesis is that there is no association between cervical squamous cell carcinoma and HIV. Mathematically stated we have; $H_o: \psi = 1$ and the alternative is; $H_1: \psi \neq 1$. Where H_0 represents the null hypothesis and H_1 represents the alternative hypothesis.

The Mantel–Haenszel procedure used in testing the statistical significance of the overall association between disease and exposure is given by :

$$\chi^2_{mh} = \frac{[\Sigma_i a_i - \Sigma_i E(a_i)]^2}{\Sigma_i V(a_i)}$$

where $E(a_i) = \frac{n_{i1} m_{i1}}{n_i}$ and $Var(a_i) = \frac{n_{i1} n_{i2} m_{i1} m_{i2}}{(n_i - 1) n_i^2}$.

Much of the statistical analysis was done using SPSS version 20.

CHAPTER 4: ANALYSIS AND RESULTS

4.0 Introduction

This chapter presents analyses and interpretation on the findings of the research.

4.1 Descriptive statistics

Table 4 below shows the distribution of the target population and number of women screened aged 18-59 years, the figures with an asterisk are projections. The researcher used 2000 and 2010 data from the Central Statistical Office to project the number of women in 2007, 2008 and 2009 using linear interpolation.

Table 4: Projection of women in the age range 18-59 and those screened.

Year	2007	2008	2009	2010	Total
Target population	2577737*	2684267*	2790797*	2897327	
Women screened	80	546	207	218	1,051

*Projected figures

Over this four-year period, there were a total of 1,051 women who were screened for cervical cancer, giving an average of 263 per year.

Out of the 1,051 screened women, the researcher took a sample of 348. For this particular sample, information was available on our variables of interest which are age, marital status, HIV status and the status of disease. The distribution of this sample over the four year period is shown in Table 5.

Table 5: Sample sizes of cases and controls by year.

Disease status	2007	2008	2009	2010
Case	13	108	29	24
Control	13	108	29	24
TOTAL	26	216	58	48

The sample sizes are not proportional to the number of women screened because data which are complete on HIV status, disease status, marital status and age were limited from year to year. For instance, the sample size in 2009 was larger than 2010 because 2009 data had more complete information than in 2010.

The sample was randomly selected using a function in SPSS and matching of age category and marital status was done for each year.

Table 6 gives a summary of a cross tabulation of age category and marital status for each year. We note that the number in each cell is either zero or an even number. The cells have even numbers because we have an equal number of cases and controls due to matching of 1 case to 1 control.

Table 6: Description of recorded variables (age and marital status) by year

		MARITAL STATUS			
YEAR	AGE CATEGORY	Single	Married	Divorced/widowed	TOTAL
	18 – 29	0	0	0	0
	30 – 39	4	4	0	8
2007	40 – 49	2	8	4	14
	50 – 59	0	2	2	4
				TOTAL	26
	18 – 29	0	0	0	0
	30 – 39	2	28	0	30
2008	40 – 49	0	116	0	116
	50 – 59	2	68	0	70
				TOTAL	216
	18 – 29	0	4	0	4
	30 – 39	0	16	2	18
2009	40 – 49	2	20	2	24
	50 – 59	0	12	0	12
				TOTAL	58
	18 – 29	0	4	0	4
	30 – 39	0	18	0	18

2010	40 – 49	0	12	0	12
	50 – 59	0	14	0	14
				TOTAL	48

In the sections that follow, we discuss the risk of cervical squamous cell carcinoma among HIV-infected women by year followed by a discussion on pooled estimates over the four year period. Testing for overall association between disease and exposure will follow and we will end the analysis by discussing the overall attributable risk.

4.2 Analyses of measure of risk for each year

In this section, we begin to discuss the measure of risk of cervical squamous cell carcinoma among HIV infected women for each year. In calculating risk, we will use exposure odds ratio (EOR) as discussed in section 3.1.2. For each year, the following will be defined;

a will represent number of cases among HIV positive women

b will represent number of controls among HIV positive women

c will represent number of cases among HIV negative women

d will represent number of controls among HIV negative women.

The estimate of EOR was given earlier and is recalled:

$$\text{Odds} = OR = \psi = \frac{Odds_1}{Odds_2} = \frac{ad}{bc},$$

The log of ψ is often used in the analysis because it has tracktable properties.

The standard error for In Ψ is given by:

$$\text{SE}(\text{In } \Psi) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

The odds ratio (OR) is used to calculate an estimate of attributable risk (AR) in a case-control study. We discussed AR in section 3.6.1 and we recall its formula below:

$$\widehat{AR} = \frac{\hat{p}(I\widehat{D}R - 1)}{\hat{p}(I\widehat{D}R - 1) + 1} \cong \frac{\hat{p}(\widehat{OR} - 1)}{\hat{p}(\widehat{OR} - 1) + 1}$$

It was noted in that section that IDR is estimated by OR and p is the estimate of proportion of HIV positive women among controls.

We used a larger data set on which information among the controls was available. For each year, the proportion (p) is obtained by dividing the number of HIV positive controls by the total number of controls. The results are shown in Table 7. The obtained estimate of p for each year was used to calculate the attributable risk for each year.

Table 7: Number of HIV positive women among controls.

Year	Number of HIV+ among controls	Total number of controls	Proportion (p)
2007	9	24	0.375
2008	10	119	0.0840
2009	26	47	0.5532
2010	10	26	0.3846

The summary of the calculations of OR and AR is shown in Table 8:

Table 8: Estimates of odds ratio and attributable risk for each year

Year	**HIV Status**	**Disease status** Case	Control	**Total**	**Odds ratio** (OR) $\hat{OR} = \frac{ad}{bc}$	**Attributable risk** (AR) $A\hat{R} = \frac{p(\hat{OR}-1)}{p(\hat{OR}-1)+1}$
2007	HIV+	9	6	15	2.625	0.3786
	HIV-	4	7	11		
2008	HIV+	14	9	23	1.6383	0.0509
	HIV-	94	99	193		
2009	HIV+	4	1	5	4.48	0.6581
	HIV-	25	28	53		
2010	HIV+	7	3	10	2.8824	0.4199
	HIV-	17	21	38		

The odds ratios show a trend upward. However, it should be noted that the years with sparse data have higher estimates of the odds ratio and attributable risk than the ones with more data. For instance, 2007 and 2010 have sparse data and have higher OR and AR than in 2008. In theory, all cells should be at least 5 in size to achieve reasonable estimates. However, all the years show estimates of odds ratio above 1. The AR also shows the same pattern as in OR. The attributable risk of cervical squamous cell carcinoma due to HIV is at least 5%. The testing of the risk of cervical squamous cell carcinoma to HIV will be done using the pooled estimates which is the subject to be discussed in the section that follows.

4.3 Pooled estimates

Due to variation from year to year seen in Table 8, aggregates estimates of OR and AR took into account this variation and the differential sample sizes.

Let $\hat{\psi}_i$ be the exposure odds ratio for year i, where i = 2007, 2008, 2009 and 2010.

Let $SE\left(ln\hat{\psi}_i\right) = \sqrt{\frac{1}{a_i}+\frac{1}{b_i}+\frac{1}{c_i}+\frac{1}{d_i}}$ be the estimates of standard error of the log of $\hat{\psi}_i$. Further, let w_i be the weight for year i in the pooled estimate of the log of $\hat{\psi}_i$. In section 3.6.2, we defined w_i to be;

$$w_i = \frac{\left(\frac{1}{SE\left(ln\hat{\psi}_i\right)}\right)^2}{\sum_{i=1}^{4}\left(\frac{1}{SE\left(ln\hat{\psi}_i\right)}\right)^2}$$

Defined in this manner ensures that the contribution of each year toward the pooled estimate is inversely proportional to the size of the standard error. The smaller the standard error the bigger the contribution. Detailed calculations of the pooled estimate of the log of exposure odds ratio are shown in the table below.

Table 9: Detailed calculations of weights

Year	$SE\left(ln\hat{\psi}_i\right)$	$\left[SE\left(ln\hat{\psi}_i\right)\right]^2$	$w_i = 1/SE\left(ln\hat{\psi}_i\right)^2$	Proportional contribution $W_i = \frac{1}{\left[SE\left(In\hat{\Psi}_i\right)\right]^2} / \sum_{i}^{n}\frac{1}{\left[SE\left(In\hat{\Psi}_i\right)\right]^2}$	$W_i ln\hat{\psi}_i$
2007	0.8189	0.6706	1.4912	0.1679	0.1620
2008	0.4509	0.2033	4.9188	0.5539	0.2735

2009	1.1514	1.3257	0.7543	0.0849	0.1273
2010	0.7633	0.5826	1.7164	0.1933	0.2046
Total				**1**	**0.7674**

We note in the table that the weight is proportional to the amount of data each year contributes. The year with most data contributed the most weight and the one with least data contributed the least. For instance, 2008 had most data and its proportional weight ($\boldsymbol{W_8 = 0.5539}$) was the highest and 2009 had the most sparse data, within the table, and contributed the least weight $(\boldsymbol{W_9 = 0.0849})$.

4.3.1 Pooled estimates of risk

The procedure of aggregating the estimates over the four year period was discussed in section 3.6.2 and we recall relevant calculations below.

Pooled estimate of ln Ψ is given by;

$$In\widehat{\Psi}_w = \widehat{\theta} = \frac{\frac{In\Psi(2007)}{[SE(In\Psi(2007))]^2} + \frac{In\Psi(2008)}{[SE(In\Psi(2008))]^2} + \frac{In\Psi(2009)}{[SE(In\Psi(2009))]^2} + \frac{In\Psi(2010)}{[SE(In\Psi(2010))]^2}}{\frac{1}{[SE(In\Psi(2007))]^2} + \frac{1}{[SE(In\Psi(2008))]^2} + \frac{1}{[SE(In\Psi(2009))]^2} + \frac{1}{[SE(In\Psi(2010))]^2}}$$

$$= W_7 \text{In } \Psi_7 + W_8 \text{In } \Psi_8 + W_9 \text{In } \Psi_9 + W_{10} \text{In } \Psi_{10}$$

$$= 0.1679 \times 0.9651 + 0.5539 \times 0.4937 + 0.0849 \times 1.4996 + 0.1933 \times 1.0586$$

$$= 0.1620 + 0.2735 + 0.1273 + 0.2046$$

= **0.7674** (Also shown in the last column of Table 9)

To obtain the pooled estimate of exposure odds ratio we take the antilog of the above expression to obtain:

$\boldsymbol{\widehat{OR} = e^{In\widehat{\Psi}_w} = 2.1542}$, as the weighted estimate of the odds ratio.

Assuming independence, the variance of the log of $\widehat{\Psi}_W$ was determined using the expression below.

$$\widehat{\text{Var}(\text{In}\widehat{\Psi}_w)} = w^2{}_7 \text{Var}(\text{In}\widehat{\Psi}_7) + w^2{}_8 \text{Var}(\text{In}\widehat{\Psi}_8) + w^2{}_9 \text{Var}(\text{In}\widehat{\Psi}_9) + w^2{}_{10} \text{Var}(\text{In}\widehat{\Psi}_{10})$$

=

$$0.1679^2\left(\frac{1}{9}+\frac{1}{6}+\frac{1}{4}+\frac{1}{7}\right)+0.5539^2\left(\frac{1}{14}+\frac{1}{9}+\frac{1}{94}+\frac{1}{99}\right)+0.0849^2\left(\frac{1}{4}+\frac{1}{1}+\frac{1}{25}+\frac{1}{28}\right)+0.1933^2\left(\frac{1}{7}+\frac{1}{3}+\frac{1}{17}+\frac{1}{21}\right)$$

$$= 0.1679^2(0.6706)+0.5539^2(0.2033)+0.0849^2(1.3257)+0.1933^2(0.5826)$$

$$= \mathbf{0.1127}$$

It can be shown that the standard error for estimate of $\widehat{\psi}_W$ is given by

$$\mathrm{SE}(\widehat{\mathrm{OR}}) = \frac{1}{\sqrt{\sum_{i=2007}^{2010} w_i}} = \frac{1}{\sqrt{1.4912+4.9188+0.7543+1.7164}} = 0.3498$$

For the overall odds ratio, the risk of having cervical squamous cell carcinoma was 2.1542 times higher for exposed females (HIV+ females) as compared to unexposed females (HIV- females) with a standared error of 0.3498. This means that the risk of developing the disease among HIV-infected women was 2.1542 times more than for HIV-uninfected women. From Table 8 above, we have shown that the odds ratio is above 1 for each year, consequently, the overall is also above 1. An approximate 95% confidence interval for ψ_W is (1.4686, 2.8398), which excludes 1, suggesting a significant difference in risk between the two groups, at the 5% level.

4.3.2 Testing for the overall association

The Mantel-Haenszel Procedure is the test that was used to assess the strength of association between the disease and exposure, to validate or invalidate the hypothesis; "The incidence rate of cervical squamous cell carcinoma is higher in HIV-infected women than HIV-uninfected women".

Table 10 shows calculations of each year which go into the Mantel-Haenszel test statistic.

The calculations are based on Table 8.

Table 10: Datailed calculations of contribution for each year toward the Mantel-Haenszel test

Year				

	ad/n	bc/n	$E(a) = n_1 m_1/n$	$V(a) = n_1 n_2 m_1 m_2/(n-1)n^2$
2007	2.4231	0.9231	7.5	1.6500
2008	6.4167	3.9167	11.5	5.1616
2009	1.9310	0.4310	2.5	1.1623
2010	3.0625	1.0625	5.0	2.0213
	13.8333	**6.3333**	**26.5**	**9.9952**

The null hypothesis is that there is no association between cervical squamous cell carcinoma and HIV.

The Mantel-Haenszel procedure utilizes the chi-square statistic with 1 degree of freedom which is given below:

$$\chi^2_{mh} = \frac{[\Sigma_i a_i - \Sigma_i E(a_i)]^2}{\Sigma_i V(a_i)}$$

$$= \frac{[34-26.5]^2}{9.9952}$$

$$= \mathbf{5.6277}$$

The value of the test for the overall association was 5.6277 and its associated P-value was 0.017679. This value suggests a significant association between cervical squamous cell carcinoma and HIV at the common significance level of 5%.

4.4 Estimate of attributable risk

Since the risk is higher in women who are HIV positive, there is need to assess the excess risk of exposure to squamous cell carcinoma. This is estimated through the attributable risk (AR) statistic. The AR was weighted over the four year period and is a function of the Odds Ratio (OR) which approximates IDR. The weighted AR is estimated by using the estimates of weights in Table 9. We recall the relevant formula in section 3.6.1.

1. $\hat{AR} = \frac{p_w(\hat{OR}_w - 1)}{p_w(\hat{OR}_w - 1)+1}$, where;
2. $\hat{OR}_w = W_7 \hat{OR}_7 + W_8 \hat{OR}_8 + W\hat{OR}_9 + W_{10}\hat{OR}_{10}$

$$\hat{OR}_w = 0.1679 \times 2.625 + 0.5539 \times 1.6383 + 0.0849 \times 4.48 + 0.1933 \times 2.8824$$

$$\hat{OR}_w = \mathbf{2.2857}$$

3. $p_w = W_7 p_7 + W_8 p_8 + W p_9 + W_{10} p_{10}$

$$p_w = 0.1679 \times 0.375 + 0.5539 \times 0.0840 + 0.0849 \times 0.5532 + 0.1933 \times 0.3846$$

$$p_w = \mathbf{0.2308}$$

Therefore;

4. $\hat{AR} = \frac{p_w(\hat{IDR}_w - 1)}{p_w(\hat{IDR}_w - 1) + 1}$

$$= \frac{0.2308(2.2857 - 1)}{0.2308(2.2857 - 1) + 1} = \mathbf{0.2288}$$

The estimate of the attributable risk associated with HIV for cervical squamous cell carcinoma was 0.2288. The interpretation is that among HIV positive, 22.88% incident cases of cervical squamous cell carcinoma are attributed to the fact that these women are HIV positive. Among HIV positive women, 22.88% incident cases of cervical squamous cell carcinoma that occur could be prevented if HIV was avoided.

4.5 Discussion

The aim of this study was to determine whether squamous cell carcinoma of the cervix does pose higher risk among HIV-infected women as compared to HIV-uninfected women and also to determine the magnitude of the risk.

This section will discuss the findings which addressed the objectives of the study and these were; (i) Assessing the statistical significance of the difference in the risk between HIV positive women and HIV negative women and (ii) Measuring the excess risk among HIV infected women if any.

In determining whether or not squamous cell carcinoma of the cervix does pose higher risk among HIV-infected women, we statistically found out that the risk of cervical squamous cell carcinoma in these women is indeed high. A weighted estimate of the exposure odds ratio of the risk of cervical squamous cell carcinoma was found to be 2.1542, suggesting that HIV positive women are twice more likely to experience cervical squamous cell carcinoma than HIV negative women. An approximate 95% confidence interval of the true odds ratio was calculated to be

(1.4686, 2.8398) which excludes 1, suggesting a significant difference in risk between the two groups at the 5% level.The data validates our hypothesis that *the incidence rate of cervical squamous cell carcinoma is higher in HIV infected women than HIV in uninfected women.*

Having determined that HIV infected women are at a higher risk of developing cervical squamous cell carcinoma, the study went on to calculate the attributable risk. The excess risk of exposure to cervical squamous cell carcinoma among HIV positive women revealed that the overall attributable risk was 0.2288, indicating that about 22.88% cases of squamous cell carcinoma among HIV positive women are attributed to the fact that these women are exposed to HIV.

Our finding of a higher risk among HIV infected women does seem to be consistent with findings of other studies reported elsewhere. For instance, Parham et al, (2006) found the prevalence of Squamous Intraepithelial Lesions (SIL) to be 56% with an odds ratio of 12.4. Although the odds ratio for Parham et al was higher, the magnitude of direction of risk is the same as in our study, they are both above 1. The difference in the values could be as a result of a difference in the research question, study design, the target age group and the sample size. Kapiga et al, (2009) showed that the prevalence of SIL in HIV-infected women was 2.9%, however, the prevalence in HIV uninfected women was not reported because their focus was on HIV infected women. Consequently, the comparison could not be made.

The difference between the findings of this study and those of other researchers may be attributed to the fact that the focus was different from our study because they were only looking at the prevalence. For instance, Kapiga looked at the prevalence and risk factors for SIL in HIV infected women while our focus was a comparative study in risk between HIV infected women and HIV uninfected women and a comparison between their study and ours could not be made. Other issues that could explain the difference include different statistical issues, among them are, target population, study design, research question and sample size. For instance, our study targeted women aged 18-59 years old while Parham et al targeted women aged 23-49 years old. In terms of study design, our study was a matched case control design which may not have been the case with other studies. The matching was done to control the influence of age and marital status. This tended to restrict our sample size in order to ensure that each case was matched to a

control on the two factors; age and marital status. In our study, the sample size was 348 while in Parham et al it was 150.This study provides the first comprehensive analysis of statistical assessment of cervical squamous cell carcinoma attributable risk to HIV infected women in Zambia and also assesses the excess risk other than determining the prevalence as was the case in other studies. The study is limited to the period 2007 to 2010.

Although the data came from all the provinces, the sample did not present each province adequately; Lusaka had a higher representation in relative terms. There is need therefore, to sample proportional number of women in each province so as to get a true reflection of the extent to which HIV positive women are at higher risk of having cervical squamous cell carcinoma. The over representation of Lusaka suggests that our study results are reflective more of Lusaka province.

However, it is clear from our findings that there is need to open more cervical cancer clinics (CCC) considering the high HIV infection rate among women who are likely to be infected with cervical squamous cell carcinoma.

CHAPTER 5: CONCLUSION AND RECOMMENDATION

5.0 General Summary

This dissertation focused on statistical asssessment of cervical squamous cell carcinoma attributable risk to HIV-infected women in Zambia. The hypothesis was that the incidence rate of cervical squamous cell carcinoma is higher in HIV-infected women than HIV-uninfected women.

The aim of the study was to determine the excess risk of squamous cell carcinoma of the cervix among HIV-infected women in Zambia. A case-control study was conducted in which secondary data was collected from National Cancer Diseases Hospital (NCDH).

5.1 Conclusion

The statistical measures used to achieve the objectives of the study showed that there is a higher risk of having squamous cell carcinoma of the cervix for women who are HIV positive compared to those who are HIV negative. The association between cervical squamous cell carcinoma and

HIV was found to be statistically significant. The higher risk of having cervical squamous cell carcinoma may be due to having limited access to the screening and treatment services offered within the country, Zambia. The results of the study indicate that there is need for appropriate implementation of screening and treatment programs for HIV-infected women in Zambia. Furthermore, if more clinics could be created for these services, then the risk of having cervical squamous cell carcinoma could be reduced.

5.2 Limitations of the study

The study has some limitations which are listed below:

- As a consequence of matching, our sample reduced in size, larger sample would have been desirable.
- Most data was from Lusaka province and the prevelence may be under estimated in the other provinces.
- The sample comprised women aged 18-59 years, hence, the results might not be generalizable to the rest of the women in Zambia.
- Hospital data often leaves out those individuals unable to access health services and may therefore, exclude a considerable number of women.

5.3 Recommendation

The researcher recommends that the ministry of health should put in place policy/programme which will provide a guidelines to HIV/AIDS caregivers and policy-makers in other organisation to ensure that they provide holistic patients' care. Such care could be optimal for reproductive health of HIV infected women if they include in their care giving cervical cancer education, routine screening and offering appropriate refferal channels for those with abnormal cells.

BIBLIOGRAPHY

Adjorlolo-Johnson1 G, Elizabeth Unger R, Boni-Ouattara E, Touré-Coulibaly K, Chantal M, Suzanne Vernon D, Marcel S, Alan Greenberg E, Stefan Wiktor Z and Terence Chorba L (2010). *Assessing the relationship between HIV infection and cervical cancer in Côte d'Ivoire: A case-control study.*

American Cancer Society (2010). *Cancer Prevention and Early Detection.* Journals.lww.com/greenjournal. National Cancer Institute.

American College of Obstetricians and Gynaecologists, (1951). *Human papillomavirus infection.*

Atashili J.,(2009). *Cervical Precancerous Lesions in HIV-infected women in Cameroon: Prevalence, Predictors and Potential impact of Screening*.

Benichou, J. (2007). *Measuring Attributabke Risk.* CR. Biologies 330, 281- 298

Beral P., Berrington V., Colin D., Franceschi S., Goodhill A., Green J., Peto J., Plummer M., Sweetland S., (2007). *Cervical Cancer and Hormonal Contraceptives*.

College of American Pathologists (2011). *Cervicle Cancer*. www.cap.org.cancerprotocols

Denny L. (2008). Prevention of cervical cancer. *Reproductive Health Matters*, 16:18-31.

Engelstad, L.P, Steward S, Otereo-Sabogal R, Leung M.S, Davis P.L, Pasick R.J. (2005,). *The effectiveness of a community outreach intervention to improve follow-up among underserved women at highest risk for cervical cancer.* Pre.Med 41(3-4): 741-748.

Feingold A. R. (1990) *"Cervical cytology abnormalities and papilloma virus virus in women infected with human immunodeficiency virus,* Journal of Acquired Immune Deficiency Syndromes.

GLOBOCAN (2008). *Cancer Incindence and Mortality Worldwide*. http://globocan.iarc.fr

Green J., Beral P., Berrington V., Colin D., Franceschi S., Goodhill A., Peto J., Plummer M., Sweetland S., (2007). *Cervical Cancer and Hormonal Contraceptives*.

James, F., (2008). *The viral origin on cervical cancer in Rabat,* International Journal on Nursing Studies, 45: 23-25, Morocco.

Kapiga S.H, Msamanga GI, Spiegelman D, Mwakyoma H, Fawzi WW, Hunter DJ (1999). *Risk factors for cervical squamous intraepithelial lesions among HIV-1 seropositive women in Dar es Salaam, Tanzania.* International Journal Gynecol Obstet 67:87–94.

Khan M.J., Partridge E.E., Wang S.S., Schiffman M. (2005). Socioeconomic status and the risk of cervical cancer intraepithelial neoplasm grade 3 among oncogenic human papillomavirus DNA-positive women with equivocal or mildly abnormal cytology cancer. 1; 104 (1): 61-70.

Kleinbaum DG, Kupper LL and Morgenstern H (1982). *Epidemiologic Research: Principles and Quantitatives Methods.* New York Willey.

National Cancer Institute, 2010.

Parham, G.P, Sahasrabuddhe, V.V., Mwanahamuntu, M.H., (2006). *Prevalence and predictors of squamous intraepithelial lesions of the cervix in HIV-infected women in Lusaka, Zambia. Gynecol Oncol,* 103:1017-1022.

Parkin, D.M. (2001). *Global cancer statistics in the year 2000*. Oncology; 2 Lancet; 533-43

Parkin, D.M. (2011). *Cancers attributable to exposure to hormones in the UK in 2010*. Britain Journer of cancer, 6; 105 (S2): S42; doi: 10. 1038/bjc. 487.

Parkin, D.M, Pisani P, Ferlay J. (1999). *Estimates of the world incidence of 25 major cancers in 1990.* International Journer of Cancer, pages 827-841.

Rabkin, C. S (1997). *"Cervical incidence trends in women at high risk of human immunodeficiency virus (HIV) infection*," International Journal of Cancer, 55 (1993),pp. 208-212, X. W.

Rosenberg, D. and Handler, A. (1998). *Population Attributable: Measures of Association and Hypotheses. HRSA*

Sahascrabuddhe W and Mwanahamuntu M.H., (2007). *Prevalence of HPV and Lesions in Zambia.*