2020

Assessing efficient odds ratios: an application to surgical stage prediction in cervical cancer.

Jesang, Jean C.
Strathmore Institute of Mathematical Sciences
Strathmore University

Assessing Efficient Risk Ratios: An Application to Surgical Stage Prediction in Cervical Cancer

Jean C. Jesang

Masters of Science in Statistical Science

2020

Assessing Efficient Odds Ratios: An Application to Surgical Stage Prediction in Cervical Cancer

Jean C. Jesang

Submitted In Partial Fulfillment of the Requirements for the Degree of Masters in Statistical Science at Strathmore University

Institute of Mathematical Sciences
Strathmore University
Nairobi, Kenya

June, 2020

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

________________________________ [Name of Candidate]

________________________________ [Signature]

________________________________ [Date]

Approval

The thesis of Jean Jesang Cheptumo was reviewed and approved by the following:

____

Dr. Collins Odhiambo
Senior Lecturer
Strathmore University

Abstract

Background: Cervical cancer remains the second most commonly diagnosed cancer and the third leading cause of cancer death in developing countries. Improving clinicians' knowledge and understanding of surgical staging is critical in the fight against the disease. Kenya has limited research on accurately predicting the surgical stage following surgical treatment for cervical cancer.The uptake of predictive mechanisms by gynecologists has not been common. **Objective**: To assess prediction by comparing the odds ratios of three popular ordinal regression models i.e. the Multinomial Logistic Regression (MLR) model, the Continuation Ratio (CR) model and Adjacent Category Logistic (ACL) model when applying cervical cancer data in surgical stage prediction. **Method**: We systematically compared the performance of MLR, CR and the ACL as the predictive mechanisms and evaluated the most appropriate model in the cervical cancer setting. The study considered women who visited the Oncology department at the Moi Teaching and Referral Hospital's Chandaria Cancer and Chronic Diseases Center and were diagnosed and surgically treated for cervical cancer from January 2014 to December 2018. **Results and conclusion**: We presented the comparison between 3 different regression models for ordinal data within the cervical cancer setting. We choose to carry out an inferential and a predictive approach. The inferential approach found that the CR model without proportional odds yielded better results when comparing the Akaike Information Criterion (AIC), log likelihood ratio and residual deviance. In addition, the key prognostic factor associated with invasive cervical cancer was the FIGO clinical stage which in particular, had a higher influence on the surgical stage 2 outcomes compared to the lesser surgical stage categories. All the 5 independent features selected for classifying the patients into surgical stages were the FIGO clinical stage and partly, the presence or absence of cancer of symptomatic vaginal discharge.
However, the predictive approach found that the MLR, CR and ACL models were not statistically different and not suitable for the prediction of the surgical stage among the women surgically treated for cervical cancer.

Keywords Surgical Stage; Ordinal Regression; Cervical Cancer; Odds Ratio; Predictive variables

Contents

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike Information Criterion
AMRS	Ampath Medical Records System
CIN	Cervical Intraepithelial lesion
CPO	Cumulative Proportional Odds
CR	Continuation Ratio model
DSS	Disease-Specific Survival
FIGO	Federation of Obstetrics and Gynecology
GLM	Generalized Linear Models
HIV	Human Immuno-deficiency Virus
HPV	Human Papilloma Virus
LEEP	Loop Electro-surgical Excision Procedure
ML	Machine Learning
MLR	Multinomial Logistic Regression
SCJ	Squamous Columnar Junction of the cervix
SIL	Squamous Intraepithelial lesion
STI	Sexually Transmitted Infections
VIA	Visual Inspection with Acetic acid

Acknowledgments

First and foremost, I would like to thank God Almighty for the strength and opportunity to pursue my Masters study.
Secondly, I take this opportunity to express my warm and heartfelt gratitude to my supervisor Dr. Odhiambo Collins for his support, guidance and patience with the invaluable constructive criticism and advice throughout the course of my Masters study. I could not have chosen a better supervisor for my Masters study.
Apart from my supervisor, I would like to thank Dr Linda Chaba (Strathmore University) and Prof. Franck Adekambi (University of Johannesburg) for their insightful comments and queries. Many thanks to the staff of Strathmore University whose administrative assistance in constant updates and communication have enabled my successful completion of my Masters study. I am very grateful to the senior lecturers who took part in the modular semesters that have eventually led to the development of this thesis.
Last but not least, I would like to acknowledge my ever supportive family Mr Joel Cheptumo, Jelagat Cheptumo, Claire Cheptumo and Charles Cheptumo who offered support and encouragement through the times spent away from home and work in pursuit of my studies and Myles Kiplimo Cheptumo, my son who was patient and understanding in times of my absence from home.

Thank you.
Miss. Jean C Jesang
Date: 26th February 2020

Chapter 1

Introduction

This chapter discusses the motivation and background of the study, research objectives and the significance of the study.

1.1 Motivation

Cervical cancer is the cancer that emanates from the cervix. The healthy normal cells on the cervix are transformed into abnormal cells that multiply and invade other parts of the body. Wilson et al. (2018) states that cervical cancer is the leading cancer in Kenya among women of all ages with a crude incidence rate of 22.4 per $100,000$ persons and a crude mortality rate of 11.5 per $100,000$ in the year 2017. Cervical cancer is caused by infection of the cervix by the human papilloma virus. Petry (2014) clearly states that the persistence of the human papilloma virus infection on the cervix causes oncogenic cell transformation at the squamous columnar junction. In their article, Wilson et al. (2018) identified HPV types 16 and 18 to be the most prevalent among women with a normal cytology, low and high grade cervical lesions and those who progress to cervical cancer. Nonetheless, cervical cancer is the best preventable malignancy of all relevant human cancers with an increase in the establishment of cervical cancer screening centers in middle and low income centers. Thus, Coleman et al. (2016) realized that the introduction of screen and treat strategies for patients with abnormal Visual Inspection with Acetic Acid (VIA) of the cervix has increased the number of women screened and treated for cervical cancer in Kenya. Howbeit, with the availability of HPV vaccines, the high costs limits their implementation in the middle and low income countries leading to more access to surgical care than chemotherapy and radiotherapy. Surgical treatment is among the curative options given to women diagnosed with cervical cancer in the middle and low income countries. The extracted specimen undergoes pathological assessment to determine the full extent of the disease thus classifying the specimen into a surgical stage.

Globally, the development and use of predictive models today is growing rapidly and highly applicable in the health care sector for the provision of efficient care and resources to patients. Moreover,James et al. (2013) defines predictive modeling as the process of developing a mathematical model for predicting or estimating an output based on one or more inputs and quantifying the accuracy of the model to make future output predictions . Predictive models are developed from statistically significant factors associated with the outcome of interest and the models can range from complex to simple. Freedman et al. (2005) reports that the National Cancer Institute had identified risk prediction as an area of extra-ordinary opportunity in the "Nation's Investment in Cancer Research". The relevance of predictive modeling today in cervical cancer care is best summarized in the words of Dr Micheal Rothberg:

"*While HPV tests are very helpful in predicting cancer risk, other factors are just as powerful at predicting cervical cancer risk. The more that we can personalize risk prediction, the more efficient our screening efforts will become(Rothberg, 2018).*"

The application of predictive modeling techniques in the early diagnosis and prognosis of cancer has become a requisite to facilitate effective clinical management of patients. Kourou et al. (2015) further support that machine learning techniques aim to model the progression and treatment outcomes of the cancer and improve our understanding of the disease thus resulting in accurate and effective management of cancer patients. The authors acknowledge that machine learning techniques could improve the accuracy of cancer susceptibility, recurrence and survival prediction. The accuracy of predictive models is clearly asserted in this statement by Kourou et al. (2015):

"*The accuracy of cancer prediction outcome has significantly improved by* 15%–20% *in the last years, with the application of ML techniques.*"

Predictive modeling methods identify patterns and their relationships in simple to complex data sets. It involves the estimation of unknown statistically significant independent variables from a particular data set and the estimated independent variables are used to predict new outputs of the outcome of interest.

According to Powers et al. (2005), predictive models can be used to risk-stratify patients and appropriately distribute resources such as caregivers and treatment combinations to the women and also, identify women who are at high risk of progression to clinical disease for disease management programs . Notably, predictive modeling in the health sector has the potential to impact clinical and therapeutic decision making.

This article gives an overview of 3 regression models developed for ranked data. It is clear that the Cumulative Proportional Odds (CPO) model is the most popular model and has naturally been utilized for analysis . However, due to the inflexibility of the proportional odds assumption, other regression models for ordinal data have been developed that ease on the proportional odds assumption. Generally, regression analysis investigates the influence of multiple predictors or independent variables on a dependent variable or outcome. The assumption of proportional odds in ordinal regression is that the effects of any explanatory variables are consistent or proportional across the different categories. One of the major shortcoming of Cumulative Proportional Odds(CPO)model is the relationship between the predictors and the response variables that can be greatly misleading when assumptions are violated.Theoretically, Wan Kai (2008) states that a more recommended model for ordinal data would take into account the categorical nature of the response since more information is contained within the ordered structure of the categories. Wan Kai (2008) further asserts that ordinal data is non-separable, independent , strictly increasing (decreasing) with arbitrary cutpoints of some underlying continuum. Based on the pathologist's point of view i.e. the surgical stage in this study, the most vital prognostic factors were presented and existing dissensions in the classification and diagnosis of the extracted tumors clarified by 3 types of regression models. In this study, we seek to assess 3 types of regression models for ordinal responses to predict the surgical stage of HIV infected and uninfected women surgically treated upon being diagnosed with cervical cancer. The 3 predictive mechanisms covered here have previously been looked by Brant (1990); Christensen (2012); McCullagh (1980); Winship and Mare (1984). Brant (1990) sought to assess the goodness of fit of the proportional odds model for ordinal logistic regression. This particular model represented a series of logistic regressions for dependent binary variables utilizing common regression parameters (with the proportional odds assumption). Winship and Mare (1984) stated that the methods for ordinal variables are considered natural extensions of probit and logit models for dichotomous variables . Winship and Mare (1984) asserts that such models explicitly recognize ordinality, avoid arbitrary assumptions concerning the ordinal scales and allow for analysis of continuous, dichotomous and ordinal variables within a common statistical framework . Statistical packages such as lme4, nnet were developed to allow for the implementation of cumulative link (mixed) models which are also known as ordered regression models, proportional odds models, proportional hazards models for group survival times and ordered logit/probit model. According to Christensen (2012), estimation techniques were mainly via maximum likelihood . Through extensions to non-linear models, McCullagh (1980) reports that the method of iteratively reweighted least squares converged to the maximum likelihood estimate which greatly simplifies the necessary computation of regression models for ordinal data. Ananth and Kleinbaum (1997) give an excellent. Ananth and Kleinbaum (1997) fully described statistical methods for modeling ordinal response data such as the continuation ratio model, the polytomous logistic model among others with application to perinatal health programme data.

1.2 Background to the study

1.2.1 Cervical cancer

According to the World Health Organization, cervical cancer is the fourth most common cancer constituting 7.9% of the female cancers. Cervical cancer affects the cervix (neck of the womb) which is the lower narrow part of the entrance to the uterus(womb). The outer surface of the cervix connects the uterus to the vagina. Figure 1.1 shows the position of the cervix in the female reproductive system.

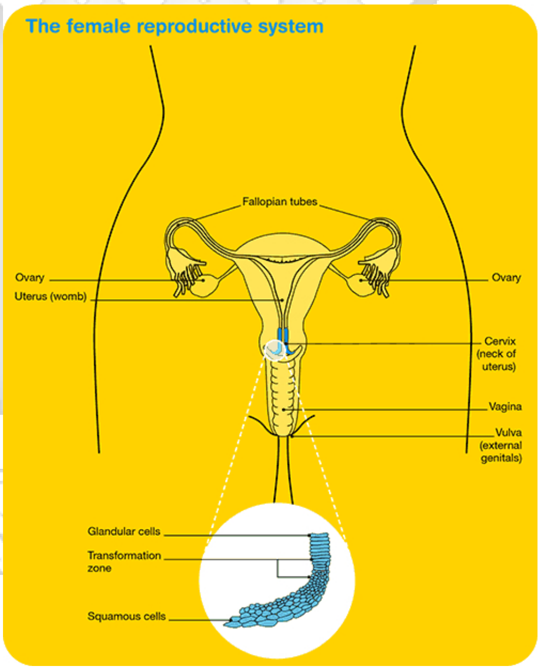

Figure 1.1: Position of the Cervix in the female reproductive system(Australia, 2015)

The surface of the cervix is covered by two types of cells: Squamous cells which are flat thin cells found on the outer part of the cervix that opens up into the vagina (ectocervix) and glandular cells which are column-shaped cells found on the inner surface of the cervix (endocervix). According to the Australian Cancer council, the area where the squamous cells and glandular cells meet is called the Squamocolumnar Junction (SCJ) and the area lateral to the SCJ is called the Transformation zone. This is where most cervical cancers start. Nordqvist (2017) explains that normal cells on the cervix have a set lifespan thus they die and are replaced by new cells.However, in some women, the normal cells in the transformation zone gradually change into abnormal cells (cells that do not die and keep dividing) which creates an area of abnormal tissue on the cervix called pre-cancerous lesions that might turn into cancer. The terms usually used to describe pre-cancerous cervical lesions include: cervical intraepithelial neoplasia (CIN), squamous intraepithelial lesion(SIL) and dysplasia. According to the American Society of Cancer (ASCO), only some women with the pre-cancerous lesions will turn into true invasive cancer. However, for most women, the pre-cancerous cells will go away without any treatment. The commonly known risk factors associated with pre-cancer of the cervix and the gradual development to invasive cancer are:

- Persistence of HPV infection which is sexually transmitted.
- Multiple sexual partners.
- Early onset of sexual activities.
- Smoking.
- A weak immune system as seen with the HIV infected women or women with immune system related diseases.
- Multiple pregnancies.
- Sexually transmitted diseases such as Chlamydia.

In an article by the International Agency for Research in Cancer (IARC), the VIA positive test outcome is termed invasive cancer of the cervix when there is a clinically visible ulcero-proliferative growth on the cervix that turns densely white after application of acetic acid and bleeds on touch. Histopathological examination of a cervical biopsy provides a definite diagnosis of invasive cervical carcinoma. In addition, imaging tests, clinical stage, type of cancer, age and parity at diagnosis are vital to the choice of treatment. The main types of invasive cancer of the cervix are squamous cell carcinoma and adenocarcinoma which are named after the types of cells that become cancerous. According to Cancer Research UK, $70 - 80\%$ of every 100 cervical cancers have the involvement of the squamous cells. Adenocarcinoma begins in the glandular cells that produce mucus which are scattered within the endocervical canal. Adenocarcinoma is less common compared to squamous cell carcinoma with 10 in every 100 cervical cancers.
Figure 1.1 clearly shows the location of squamous cell carcinoma development in the skin-like cells that cover the outside of the cervix and adenocarcinoma development in the glandular cells. In addition, the determination of the clinical stage of cervical cancer depends on the size of the tumor. However, the invasive physical check up for the tumor size and extraction of a biopsy for

histopathological evaluation takes a trained medical professional. This is labor intensive, time consuming and prone to errors.

Figure 1.2 shows the position of the SCJ on the cervix.

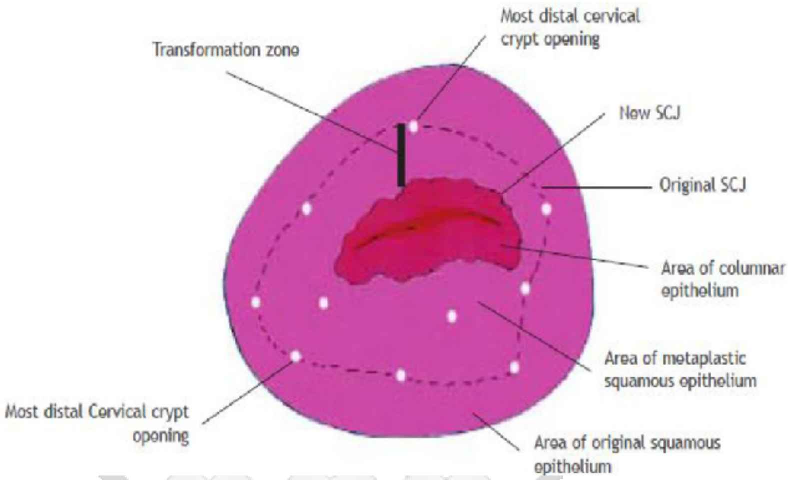

Figure 1.2: Position of the SCJ on the Cervix(Pradjatmo, 2015)

According to Atashili (2009), the psychological changes in the size and anatomical position of the cervical SCJ may play a role in the age at which lesions occur. In addition, he noted that the age at which lesions occur may also be related to the age specific incidence, prevalence and persistence of HPV infection. Therefore, the incidence of and mortality due to cervical cancer increases with age. Additionally, in an article by Sangwa-Lugoma et al. (2011), the presence of high risk HPV is associated with the woman's behavioral and sexual characteristics which affect the risk of new infections and immune function. Moreover, Chambuso et al. (2016) states that the presence of HIV is highly significant to the progression of a HPV infected woman to clinical disease due to the higher chances of recurrent HPV infections. Abraham et al. (2013) discovered that the incidence of cervical cancer among HIV infected women increased consistently with diminishing $CD4^{+}$ count leading to reduced immune function thus it is strongly evident that there is a biologic relationship between HIV status and clinical disease outcome. In contrast, since the introduction of antiretroviral therapy and the triple drug combinations, HIV cannot be assumed to be solely involved in the oncogenic HPV progression to clinical disease. Moreover, Denslow et al. (2014) discovered a gap in existing literature on the effect of antiretroviral therapy on the progression to clinical disease. Also, sexually transmitted diseases have been found by Sonia et al. (2017) to be associated with high risk HPV infection and persistence with bacterial vaginosis being a highly prevalent risk factor for acquisition and transmission of many STI's, including HIV and HPV.

The prevention and treatment of cervical cancer among HIV infected women is described as complicated due to the weakened immune system increasing the risk of developing the disease. Moreover, other factors such as the history of sexual activity, age at sex debut, lifetime sexual partners, sexually transmitted disease history, prior cervical screening and the socio-demographic characteristics of the women increase the complications of oncogenic HPV progression to clinical disease. While the awareness and uptake of anti-retro viral medication has greatly improved in Kenya, there seems to be increasing cases of cervical cancer among the HIV infected women.

1.2.2 Stages of Cervical Cancer

According to the American Cancer Society 2018, clinical staging gives an estimate of the extent of the cancer. Clinical staging is done on initial diagnosis and prior to any treatment plan being assigned. It is based on a visual and physical examination of the cervix by an experienced gynecologist, imaging tests such as MRI, CT and PET scans, cervical histology results and blood tests. Heaps and Berek (1990) state that CT scans are perhaps the best technique for cervical cancer evaluation as it can assess the primary tumor, urinary and gastrointestinal tracts, liver parenchyma and the retroperitoneum. A clinical stage of a cancer does not change overtime. For instance, the information that would be added to the initial clinical stage if the cancer spreads to the bone is bone metastasis. Clinical staging is considered vital to aiding in treatment plan decision making and as a baseline when assessing how the cancer responds to treatment. Moi Teaching and Referral's Chandaria Cancer and Chronic Diseases Center use the FIGO staging system. The overall clinical stage of a cervical cancer is determined by the primary tumor size, the grade of the cancer, lymph node involvement and metastasis of the cancer. The grade refers to the differentiation of the cells of the cervix. Low grade or well differentiated cervical cancer displays cervical cells that look like normal cells from a normal tissue and the cancer gradually grows. High grade or poorly differentiated cervical cancer displays cells that appear very different from normal cells, grow quick and give a worse outcome. It is stated that whenever the grade does not affect the cancer's stage, it may still affect the outlook and / or the treatment plan. Moreover, Heaps and Berek (1990) state that CT scans are perhaps the best technique for cervical cancer evaluation as it can assess the primary tumor, urinary and gastrointestinal tracts, liver parenchyma and the retroperitoneum. However, this study is limited to recorded clinical data since the molecular scale data based on imaging is expensive and unavailable for analysis.

Figure 1.3 clearly shows the spread of the cancer from the cervix as the tumor increases in size.

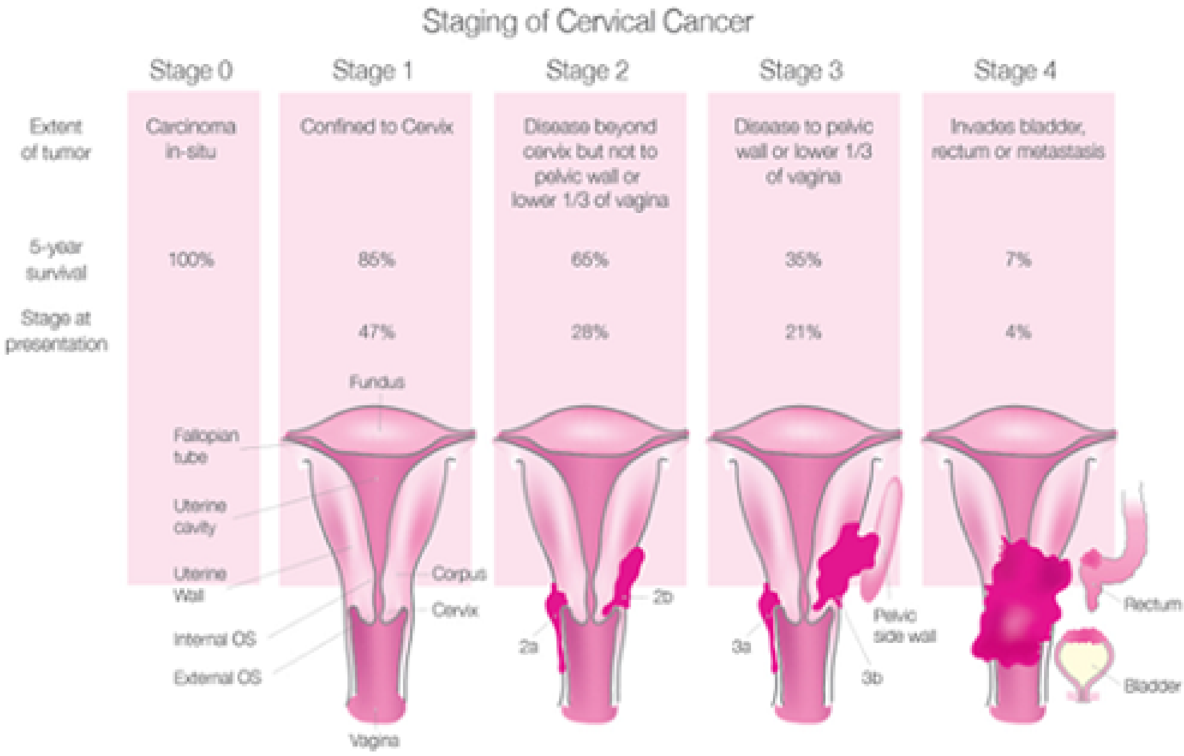

Figure 1.3: Pictorial presentation of the various stages of cervical cancer(Vinci, 2014)

The early cervical cancer tumors which are confined to the cervix include carcinoma in situ (stage 0), stage 1A , stage 1A1 , stage 1A2 , stage 1B and stage 1B1 tumors. The locally advanced cervical cancer tumors are stage 1B2 , stage 2 , stage 2A , stage 2B , stage 3 , stage 3A , stage 3B and stage 4A with stage 4B tumor being advanced cervical cancer which has spread to organs away from the cervix such as the lungs.

1.2.3 Treatment of Invasive Cancer

The main treatments offered at the Chandaria Cancer and Chronic diseases Center are surgery and chemotherapy with referrals being offered for radiotherapy treatment. Surgery involves the removal of the tumor and the tissues or organs around the tumor. The main surgical treatments given are cone biopsy which involves removal of a cone-shaped tissue of the tumor and a small area of the surrounding tissue, simple hysterectomy which involves the removal of the cervix and some pelvic lymph nodes, radical hysterectomy which involves the removal of the cervix, womb, fallopian tubes, the parametrium, the top of the vagina and lymph nodes and radical trachelectomy which is the removal of the cervix, parametrium, upper part of the vagina, the pelvic lymph nodes with the womb and ovaries left intact to enable the woman to conceive.

The International Agency for Research in Cancer explains that radical radiotherapy involves the use of x-rays to kill the cancerous cells. External radiotherapy is administered from outside the cervix and brachytherapy or internal radiotherapy is administered from inside the cervix. Moreover, chemotherapy involves the use of drugs given through a drip into the vein to kill the cancerous cells. Chemotherapy is classified as either neoadjuvant chemotherapy or adjuvant chemotherapy. If the purpose of the treatment is to shrink the tumor to ease removal or increase the effectiveness of radiotherapy, then neoadjuvant chemotherapy treatment is given before surgery or with radiotherapy. If the purpose of the treatment is to kill any remaining cancer cells after surgery or radiation, then adjuvant chemotherapy treatment is given. Furthermore, chemoradiation treatment can be given which involves using a combination of chemotherapy and radiation.

Surgical treatment options at the Moi Teaching and Referral's Chandaria Cancer and Chronic diseases Center are given to women who present clinical stages $1A - 2A1$ inclusive of carcinoma in situ. The recommended surgical treatment options for invasive cervical carcinoma are LEEP, cone biopsy, radical trachelectomy (to maintain fertility), simple (total) hysterectomy and radical hysterectomy. Upon completion of a surgical procedure, the extracted tissue undergoes surgical staging to fully assess the extent of the disease. This is a vital process as it allows for informed decision making in regards to additional treatment options such as chemotherapy, radiation or a combination of treatments. Moreover, surgical staging gives precise information on treatment response and prognosis. The surgical staging involves determining the size of the primary tumor, its location, assessing metastasis into the lymph nodes and assessing for other tumors. Besides, Matsuo et al. (2015) states that depending on the pattern of surgical risk factors obtained from the surgical specimen, patients with an increased risk of recurrence are assigned post-operative therapy. For the surgically treated patients, Matsuo et al. (2015) lists the high risk surgico-pathological factors to include pelvic lymph node metastasis, parametrial involvement and positive surgical margins whereas the intermediate risk factors include large tumor size, lymphovascular space invasion, and deep cervical stromal invasion. Pecorelli and Odicino (2003) considered that surgical and pathological data are vital for precise analysis of survival and prognostic risk factors. Thus, the surgical stage provides additional and clearer information to determine a woman's prognosis. The aim of the study was to evaluate the significance of the surgical stage to the choice by the gynecologist to surgically treat a patient based on the clinical stage and key predictive factors at the MTRH's CCCDC Oncology department. Figure

1.4 is the conceptual framework for this study which is informed by Drolet et al. (2013) who were interested in the risk of cervical cancer based on the inequalities of women in different socio-economical groups and sexual behaviors.

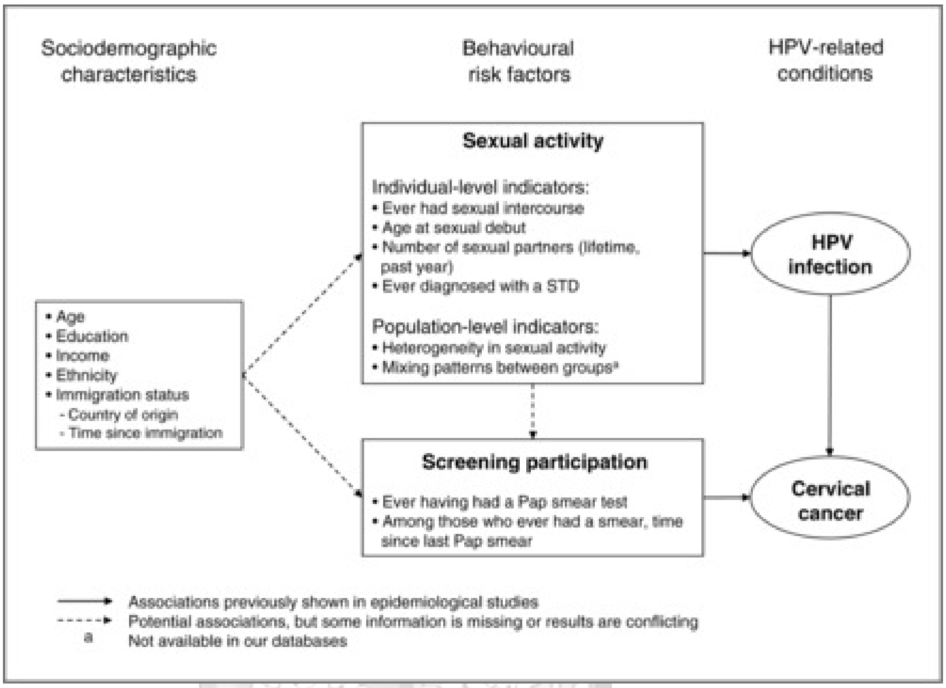

Figure 1.4: Conceptual framework of the different pathways linking sociodemographic characteristics and HPV-related conditions(Drolet et al., 2013).

Currently, clinical prediction to diagnose and treat cervical cancer is mainly done through observation, intuition, experience and wisdom. The status of a patient following the prescribed treatment is always associated with the treatment. In contrast,Rosenfeld and Breslow (2008) found that treatments that were ineffective or deleterious were perpetuated because there was no basis for determining whether there was a cause and effect relationship between the treatment and the result. It has become vital to better understand the causal connection between treatment and the clinical outcome so as to accurately assess the improving or deteriorating status of a patient on or after treatment. By utilizing inferential techniques, support systems have been created to assist doctors and clinicians detect associations in medical data and observe patterns that may be life threatening to patients. However, effective treatment of cervical cancer among the HIV infected and uninfected women in a predictive manner and the ability to automate the best treatment plan warrants investigation in Kenya.

1.2.4 Machine Learning

James et al. (2013) define machine learning as an extensive set of tools for understanding data. The tools are categorized as supervised learning and unsupervised learning tools. This study utilized the supervised learning tools which involve building a statistical model for predicting an outcome based on one or several predictor variables. The machine learning problem in this study is a classification problem. That is, the predicted outcome is a qualitative or non-numeric value. Classification is synonymous with pattern recognition. Classification involves identifying the class or category to which a new observation belongs, based on a training set of observations whose class membership is already known. Classification can be categorized into binary classification and multiclass classification. Binary classification involves only two possible classes whereas multiclass classification involves three or more classes. Machine learning algorithms perform predictions and learn from large sets of data. An algorithm that implements classification is referred to as a classifier and assigns a clinical outcome based on the observed attributes. The performance of a classifier is evaluated by various factual tests. Precision is the conventionally used performance metric test to determine the quality of a classifier. Machine learning is an extension of predictive modeling. Thus, in this study, machine learning was utilized to predict the surgical stage of the cervical disease after surgical treatment.

Predictive Modeling

The model of choice should accurately predict the expected surgical outcome following surgery. However, interest lies in the accuracy of the predictions that are obtained when the machine learning tool is applied to unobserved data. Using data from the past four years of diagnosed cervical cancer patients, we can train the machine learning method (algorithm) to predict the surgical stage following surgical treatment based on clinical measurements. Cousins et al. (2002) states that a classifier will enable women to be risk-stratified to identify opportunities for intervention prior to the occurrence of adverse outcomes thus leading to efficient resource allocation with cost-effectiveness and successful treatment of cervical cancer. In the health sector, predictive models are applied to patient medical data and the outcomes utilized to help with decision making on interventions such as drugs, physicians, treatment plans among others.

Cancer care revolves around prediction of the prognosis prior to treatment. In their article, Seel et al. (2012) state that empiricaly based, probabilistic models were more accurate than the subjective clinical judgements in predicting health outcomes and also, by generating empirical, quantitative estimates of the likelihood of an outcome, prediction models can improve accuracy. According to Vickers (2011), prediction models estimate the probability of future events. The absence of symptoms in early stage cervical cancer patients has limited treatment and recommendations to be based on the fact that tumor progression would eventually threaten a patient's morbidity. Vickers (2011) affirms that the establishment of statistical tools has provided a quantifiable probability estimate of an outcome for each patient. Moreover, the author asserted that prediction models have greater accuracy compared to clinical staging or risk grouping, integrate novel predictors and enable intelligent informed - based treatment decisions. He concluded that prediction models should be used in cancer care decision making with integration into the electronic health records and careful evaluation of model's clinical outcomes.

While it is clear that predictive models are vital in determining treatment plans for patients with chronic diseases, Rosenfeld and Breslow (2008) states that
"'it would be desirable to monitor the progress of the patient using predictive models to evaluate the effectiveness of the treatment plan on a continuous basis and to revise the treatment plan and /or the patient specific rules accordingly."

For instance, Biewenga et al. (2011) was able to show that disease-specific survival (DSS) in patients with early stage cervical cancer can be predicted with a statistical model. Upon the correction of the 9% over fit shown by bootstrap internal validation, the formula for the multivariate prognostic model to calculate the DSS at a particular time x is shown below:

$$
\begin{aligned}
&5-(year)DSS(tx)\\
&= basicsurvival(tx)^{e}xp(\beta_0(tx)+0.91((0.02*TD)+(1.6*AC)+(0.46*ASC)+(1.2*LNM)+(0.24*DI)+(0.85*LVSI)+(1.08*PMI)))
\end{aligned}
\tag{1.1}
$$

The accuracy of this particular model to make predictions was done by comparing the predicted and observed DSS at 5 year follow up. The area under the receiver operating characteristic (ROC) curve, the sensitivity and specificity were calculated to assess the ability of the model to correctly classify the patients with DSS. The authors found that after correction by a shrink factor of 0.91, the independent and statistically significant predictors used in the multivariate model were the tumor diameter (TD), histological type (either AC for adenocarcinoma or ASC for adenosquamous cell carcinoma), presence of lymph node metastasis (LNM), depth of stromal invasion in millimeters (DM), lymph vascular space invasion (LVSI) and parametrial invasion (PMI). Another example of a predictive model is by Yuri et al. (2016) who designed a prostate cancer risk calculator to predict the risk of prostate cancer by use of chi-square analysis,Kolmogorov-Smirnov test, multiple logistic regressions and the receiver operating characteristic (ROC) curve. The authors found that age, prostate volume, serum prostate specific antigen (PSA), digital rectum examination finding and family history are factors associated prostate cancer.The four parameters associated with the predictors PSA, age, prostate volume and digital rectum examination are Lpsa which is the $\log_2$(PSA), lpsac the mean value of the $\log_2$(PSA), Lvol the $\log_2$(prostate volume) and lvolc the mean value of $\log_2$(vol) The predictive equation or logistic model developed is shown below:

$$\text{Ln(odds)} = \beta_0 + \beta_1(lpsa - lpsac) + \beta_2(lvol - lvolc) + \beta_3(lage - lagec) + \beta_4(DRE) + \beta_5(HYS) \tag{1.2}$$

1.3 Problem statement

We aspire for all the women attending the Oncology department at the Chandaria Cancer and Chronic Diseases center for cervical cancer treatment be given the most effective treatment upon confirmed diagnosis for the best possible outcomes. Today, there are women with cervical cancer who have their treatment regimen changed while undergoing treatment due to the ineffectiveness of the initial regimen. This problem leads to wastage of valuable patient time and resources resulting in costly treatment and low survival rates. Therefore, we have described, illustrated and compared the odds ratios of 3 regression models for ordinal responses to predict the surgical stage of HIV infected and uninfected women surgically treated for cervical cancer. Furthermore, we seek to determine the most appropriate model for the cervical cancer data at the Oncology department.

1.4 Objectives of the Study

The general objective of the study is to assess efficient risk ratios by predicting the surgical stage outcome of women surgically treated for cervical cancer. The specific objective of the study

1. To develop, validate and compare the odds ratio of 3 types of regression models for ordinal responses which include the Continuation Ratio model, the Adjacent Category model and the Polytomous (Multinomial) logistic model.

1.5 Research Questions

The surgical outcomes in this study are surgical stage 0, surgical stage 1, surgical stage 2 and surgical stage 4. The ultimate objective is to assess prediction in terms of the odds ratio by answering the question: Among the three regression models for ordinal responses, which model best estimates and predicts the surgical stage prior to surgical treatment of cervical cancer?

1.6 Scope of the study

The main focus of this study was to develop, validate and compare the odds ratios of 3 types of regression models for ordinal responses to predict the surgical stage of HIV infected and uninfected women surgically treated for cervical cancer in the Oncology department at the Chandaria Cancer and Chronic diseases center based in Eldoret, Kenya. Retrospective complete medical records collected from women who had been diagnosed with oncogenic HPV infection and had progressed to clinical cervical cancer disease from January 2014 to December 2018 were retrieved and analyzed to develop the predictive models. This particular time period was chosen as the data was readily available. The explanatory variables considered were the socio-demographic characteristics such as age, bmi, occupation, parity,sexual behaviors such as use of contraception, clinical and pathological variables such as symptoms, method of detection, confirmed histology result, clinical stage, tumor size or grade, lymph node metastases, parametrial involvement, treatment plan and follow up visits information which include uptake of radiation, chemotherapy and recurrence of disease.

Limitations of the Study

To begin with, the study is limited by the use of an existing data set with a small sample of women who underwent surgical treatment and the utilization of 3 response categories which is considered the simplest extension of the binary logistic model . The data utilized was not specifically collected to develop predictive models or for predictive purposes. Secondly, the quality of the data was briefly addressed through data cleaning. Thirdly, the presence of bias was due to the fact that the available data was based on human memory and human error. Finally, the data may not reflect the actual extent of the prevalence of cervical cancer within Kenya since there are other fully operational institutions that offer cervical cancer treatment within and without Eldoret town. In conclusion, further research should be carried out in the future with a larger sample size to allow for cross validation techniques.

1.7 Significance of the study

The study will bring awareness to this particular oncology team on the possible key predictors and informed decision making of surgical treatment candidacy leading to proper and specific types of information being collected from patients with invasive cervical cancer. Secondly, encourage more research on algorithms that can aid with future prediction on possible outcomes thus enabling better utilization of the available resources to maximize patient benefits whilst minimizing the risk. Thirdly, the results from the study will create opportunities for the establishment of low cost modalities for evaluation of disease extent to reduce morbidity and save on resources such as pre-treatment surgical staging techniques. Fourthly, the study outcomes will increase early detection of cervical cancer by opening of fully operational national cervical cancer screening clinics with highly trained doctors, clinicians, nurses , variety of up to date screening techniques (for example pap tests, visual inspection methods, liquid based cytology, DNA HPV tests, laparoscopy and lymphadenectomy among others) and treatment facilities offering affordable surgical treatment.

Chapter 2

Literature Review

2.1 Introduction

Although surgical staging has provided data on the frequency of para-aortic lymph node metastasis by stage of cervical cancer and led to better treatment strategies, Heaps and Berek (1990) report that high rates of pelvic and systemic failure during surgical staging has produced an unexceptional boost in survival rates. However, in comparison to surgical staging, Gold et al. (2008) report that clinical staging is 60% accurate. Currently, cervical cancer remains the only gynecologic malignancy that is staged clinically. Gold et al. (2008) report that the clinical staging errors are associated with misdiagnosed lymph node metastasis. In order to estimate the absolute probability of risk for specific individual outcomes, studies have shown that different types of information on an individual at a given time can be utilized which leads to the establishment of prognostic or prediction models. Huang et al. (2010) mention that mathematical modeling applications can improve outcome predictions of cancer therapy. In the healthcare environment, predicting a condition based on clinical characteristics, observation and wisdom forms the basis of treatment given to cancer patients. However, increasing technological advancement, improvement in global health care and the diverse treatments given to cancer patients have led to the need for doctors and clinicians to recognize and deal with the challenges of effective diagnosis and treatment of cancer. The following review of literature is to determine the usability of predictive tools prior to surgical treatment of cervical cancer in the oncology health care system.

2.2 Defining the problem

According to the World Health Organization, cervical cancer screening aims to detect pre-cancerous changes which, if not treated, may lead to cancer. Additionally, it was noted that the availability of effective treatment among women $30 - 49$ years of age is not being captured. Based on the 2030 Agenda for Sustainable Development and its 17 Sustainable Development Goals, good health and well-being is one of the goals to transform our world. The Gynecologic Oncology Group assert that surgical staging has been a required method of aortic lymph node evaluation. During surgery, the gynecologists decide on the best treatment for the patient upon examining the actual extent of the disease. Furthermore, the cancer is categorized into stage $1, 2, 3$ or 4. Bhatla et al. (2019) consider pathological confirmation as the gold standard for surgicopathological assessment of lymph node involvement in surgically treated patients. At most, surgically staging for cervical cancer patients is done prior to therapy assignment through laparotomy or laparoscopy surgical procedures.

Marnitz et al. (2005) state that the FIGO clinical staging of cervical cancer is associated with a high percentage of over and under staging the tumor extent. The authors discovered that laparoscopy transperitoneal staging procedure provided more information on the actual extent of cervical cancer disease prior to surgery. Marnitz et al. (2005) studied 84 patients with cervical cancer FIGO stage I to stage IV and found that the clinical stages were underestimated for 87% of the cases with the distribution of the clinical stages according to FIGO as follows: stage IB1 in 15.5%, IB2 in 15.5%, IIA in 8.3%, IIB in 23.8%, IIIA in 8.3%,IIIB in 21.4%, IVA in 6%, and IVB in 1.2%. Pathological findings following laparoscopic staging in 43% of the patients had a higher tumor stage diagnosed. However, the laparoscopic approach is expensive and not readily available at the Oncology department at the MTRH's Chandaria Cancer and Chronic Diseases Center.

In addition, Prasad et al. (2014) postulate that the major limitations of FIGO clinical staging of cervical cancer are the discrepancies with surgical staging and the inability to assess the lymph nodes. Moreover, the authors state that the FIGO clinical staging system results in under or over-staging in $20 - 40\%$ of patients when compared with surgical staging. Prasad et al. (2014) carried out a literature review and found that CT imaging does indeed improve the accuracy of clinical staging performed by FIGO guidelines. In their study, Prasad et al. (2014) sought to determine the differences between clinical and CT staging based on data from 53 patients with cervical carcinoma prior to any treatment. A McNemar test determined that the difference between FIGO and CT staging was significant ($P < 0.001$) . Moreover, a Kappa value of 0.08 ± 0.06 suggested a poor consistency between the two staging techniques. The CT scans were visible in 70% of the patients. Vaginal involvement was seen in 61.5% of the patients during clinical examination with 41.5% being consistent in the evaluation of parametrial invasion. Additionally, unilateral pelvic involvement, bilateral involvement, pelvic nodes involvement and retroperitoneal nodes were seen in 26.4%, 32%, 20.7% and 11.3% of patients. Thus, CT imaging being the less cheaper and widely available option in developing countries, Prasad et al. (2014) findings support that CT imaging does not reliably correlate with clinical FIGO staging of cervical cancer. Nonetheless, the MRI/CT scans tend to be costly for most women attending the oncology treatment clinics particularly in the Kenyan health sector. At most, the patients are required to have CT imaging upon histological diagnosis of the cancer and after surgery to determine the need for further therapy. Moreover, the records of the CT images are not kept in the oncology databases.

Surgical evaluation allows for the stratification of patients based on the extent of the cervical carcinoma. Fader and Rose (2007) state that a surgical procedure is primarily for the establishment of a diagnosis, determine the disease extent and remove as much of the gross tumor. Averette et al. (1972) evaluated the laparotomy procedure for surgical staging prior to cervical cancer treatment. The authors discovered that clinical staging was not accurate in 38.6% of the patients and 14.3% had aortic lymph node metastasis. Therefore, they concluded that surgical staging greatly modified therapy whether primary therapy was by radical operation or radiology. Moreover, the purpose of staging according to Derks et al. (2016) is to assess the surgical treatment outcome in terms of patient recurrence patterns and survival. By applying the Kaplan-Meier technique on 129 women treated by radical hysterectomy with pelvic lymphadenectomy for stage 1B2 / 2A2, Derks et al. (2016) found a disease-specific survival of 84%. Also, by carrying out univariate and multivariate analysis, the authors realized that the histologic type, tumor diameter and the parametrial involvement were key prognostic factors associated with patient survival. Thus, 50% received adjuvant treatment and a pelvic recurrence rate of 8% was noted.

Additionally, Rutledge et al. (2004) sought to define predictive factors of nodal status and recurrence in 109 and 86 patients with stage 1B1 and 1B2 cervical cancer respectively treated with radical hysterectomy. Multivariate analysis via stepwise logistic regression identified lymphovascular space involvement (OR 6.4, CI $2.4-17$, P $=$ 0.0002) and parametrial involvement (OR 8, CI $3.1-20$, P $=$ 0.0001) as independent predictors of positive nodes. Kaplan-Meier technique was applied to evaluate the Disease Free-survival post-surgery and established that lymphovascular space involvement (ROR 5.7, CI $2-16$, P $= 0.0064$) and outer 2/3 depth of invasion (OR 5.8, CI $2-20$, P $= 0.0029$) were the independent predictors. Based on their analysis, Rutledge et al. (2004) deduced that these factors were best determined post-surgery and that treatment decisions based on tumor size only should be reconsidered.

In support of Rutledge et al. (2004), Qin et al. (2009) inferred that for operable cervical cancer, the clinical stage on its own is not well grounded for the selection of post operative therapies. Since cervical cancer is the only gynecological cancer that is staged clinically and not pathologically, Qin et al. (2009) sought to analyze

the discrepancies between the clinical stage (accuracy of the physical pelvic examination) and the pathological result. They analyzed 818 women with confirmed stage 1B to 2B cervical carcinoma treated primarily with surgery. The women were assigned to pT category based on the pathological findings. The authors noted absolute consistency of 53.1% between the clinical stage and pT category for stage 1B to 2B with a 37.3% over estimation and 9.7% underestimation. The consistency in stage 1B1, 1B2, 2A and 2B were found to be 85.4%, 77.4%, 35.3% and 20.5% respectively. In contrast, they found that the most significant inconsistency was in the accuracy of clinical staging to determine vaginal and parametrial disease which was at 70.2% and 74.0% respectively. In their conclusion, Qin et al. (2009) suggested that a surgical staging system ought to be put into consideration for operable cervical cancer tumors. Earlier on, van Nagell Jr et al. (1971) expressed that the most common error was failure to accurately define the extent of parametrial disease by clinically staging cervical carcinoma patients. van Nagell Jr et al. (1971) had retrospectively investigated the inevitable discrepancies between clinical staging and pathological findings in 370 women who were evaluated, staged and treated for invasive cervical carcinoma. There was decreasing clinical stage accuracy for stage 1 to stage 2 at 78% to 25% with only 66% of women found to have been clinically staged accurately.
While the clinical stage of cervical carcinoma is very helpful in predicting the tumor size, the vaginal and parametrial involvement, surgical staging can provide further vital information that can impact therapy. Sayed (2019) wrote that the next step in pathology is machine learning. At the level of cancer diagnosis, pathologists are accurate; however, when predicting the development of cancer, Sayed (2019) reports that the accuracy rate is at 60%. Currently, machine learning is headed towards developing a high accuracy pathology system in cancer care.Although Kenya is far from attaining low-cost technology for microarray expression, Chen (2012) carried out a study whereby he compared ordinal and non-ordinal algorithms on 5 cancer stage microarray datasets. By utilizing only the predictive genes to model the cancer stage, the author concluded that a strict ordinal classifier trained by a validated approach can more accurately predict the cancer stage compared to non-ordinal classifiers without considering the order of the stages.

2.3 Searching for solutions

The surgical stage has been found to clearly differ from the determined clinical stage at diagnosis. For instance, when surgery shows that the cancer has spread more than was perceived during the preoperative staging. According to Lagasse et al. (1980), the Gynecologic Oncology Group (GOG) confirmed that clinical staging is

often inaccurate in determining the extent of disease in cervical cancer patients. The authors arrived at this denouement in a multi-institutional study where 29% of patients with stage $2, 3$ and 4 had paraaortic nodal metastases which did not correlate with the histologic grade of the tumors. Chow et al. (2008) states that a predictive model has the potential to lead to improvement in delivery of care and to minimize under treatment or over treatment. Based on 207 stage 1B - 2A cervical cancer patients, Sun et al. (2011) developed a multivariate logistic regression model to predict lymph node metastasis. The logit model 2.1 was a proposed model to improve conservative surgery for patients with early stage cervical cancer.
Equation 2.1 was the model that was developed with a predictive accuracy of 76.3%:

$$\text{Logit}(p) = -2.534 + \text{SerumSCCAg} * 1.934 + \text{Depth of cervical stroma} * 0.801 \qquad (2.1)$$

The SerumSCCAg and depth of cervical stroma were found to be the independent risk factors of pelvic lymph node metastasis in early stage cervical cancer.
Under surgical treatment, radical hysterectomy with parametrectomy is the treatment for early stage cervical cancer. Hsu et al. (2018) report that even though the prognosis is good, parametrectomy damages the autonomic nerve and bladder functions leading to a poor quality of life. In order to improve the quality of life for eligible radical hysterectomy patients, Hsu et al. (2018) carried out a retrospective study on 339 women with class $II - III$ radical hysterectomy to develop an algorithm based on the factors predictive for possible parametrial invasion in early cervical cancer. The authors aim was to facilitate radical hysterectomy patient selection among those who have been recommended for parametrectomy. The authors applied the non - parametric chi-square test, the Krustal - Wallis test and the Mann - Whitney U test on the categorical and continuous variables. They proposed two frameworks based on either the frozen pathologic sections of the pelvic lymph nodes or the patient's age to determine whether parametrectomy is recommended.
Figure 2.3 recommends sentinel lymph node sampling with frozen pathology prior to surgery. There are higher incidences of parametrial invasion among patients with lymph node metastasis (25%(4/16) stage $1B1$,$p < 0.001$, chi - square test) and a tumor size $\leq$ 2cm.

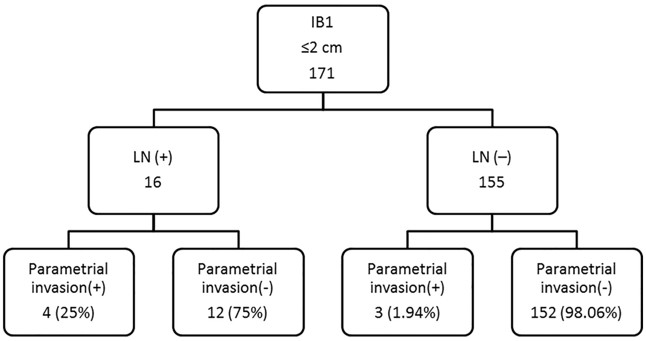

Figure 2.1: Determination of the need for parametrectomy during radical hysterectomy based on the status of pelvic lymph node metastasis (Hsu et al., 2018).

Figure 2.3 recommends radical hysterectomy prior to surgery for all patients aged $\geq$ 50 in facilities that without resources to perform sentinel lymph node sampling. The authors found that 9.7%(6/62) of the patients aged over 50 years had parametrial invasion.

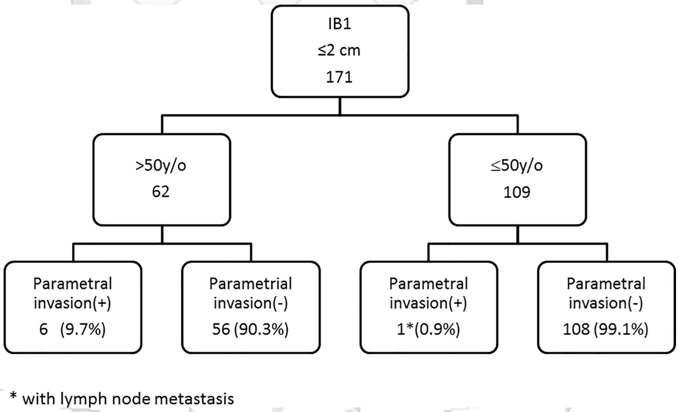

Figure 2.2: Determination of the need for radical parametrectomy during hysterectomy based on patient age.

Pathological evaluation post surgery allows for optimal benefit and minimal risk in terms of patient survival and disease recurrence. Kamura et al. (1992) maintain that based on prognostic findings, the prediction of post surgery prognosis is more precise compared to clinical staging. Kamura et al. (1992) applied multivariate analysis using cox regression models on 345 patients with stage $IB-III$ cervical cancer who had undergone radical hysterectomy and pelvic lymphadenectomy. The authors discovered that the key independent and prognostic factors on the survival of the patient post surgery were the histologic subtype, longitudinal tumor diameter and pelvic lymph node metastases. Furthermore, Gold et al. (2008) multivariate analysis on 555 surgically sampled and 130 radically evaluated using CT, MRI or lymphangiogram for para-aortic lymph node metastasis. Gold et al. (2008) established that there was poorer prognosis on the radically evaluated patients compared to the surgically staged patients. In this case, the multivariate analysis was adjusted for age, race, performance status, histology, tumor size and extent of parametrial involvement. The authors found that the relative risk of disease progression in the radically evaluated compared to the surgically staged was $1.35(95\%CI[1.01 - 1.81], P = 0.043)$ and the relative risk of death was $1.46(95\%CI[1.08 - 1.99], P = 0.014)$.

Predictive models have been successfully applied to other cancerous diseases such as prostate cancer. Partin et al. (1997) developed a multi-institutional model based on 4133 men treated by radical retropubic prostatectomy for clinically localized prostate cancer. The authors carried out a multinomial log-linear regression to concurrently predict organ-confined disease, isolated capsular penetration, seminal vesicle involvement and pelvic lymph node involvement. Then, bootstrap estimates of the predicted probabilities were used to develop nomograms to predict the pathological stage of the disease. The performance of the nomograms were validated by applying bootstrap analysis. Prostate-specific antigen level, TNM clinical stage and the Gleason score were found to be statistically significant factors for the prediction of the pathological stage with P-values< 0.001. The validation pathological stage outcomes of the predictive model were 67.3% organ-confined disease, 59.6% isolated capsular penetration, 79.6% seminal vesicle involvement and 82.9% pelvic lymph node involvement. Thus, Partin et al. (1997) proposed that the nomograms would enable physicians and patients to make more informed based decisions and individual patient counseling based on the probability of the tumor being at a particular stage.In support of predictive models, Gancarczyk et al. (2003) developed probability nomograms for pathological stage outcome at the time of radical prostatectomy for $1,510$ men. Univariate and multivariate logistic regression analysis established that pretreatment PSA, highest biopsy Gleason sum and the percentage of cores positive for cancer were the most significant independent predictors of pathological stage. Moreover, Partin tables have been developed to predict the pathological stage at radical prostatectomy. A table for clinical unilateral T3a prostate cancer patients was developed by Joniau et al. (2007) based on preoperative prognostic factors of 200 men with the disease who underwent a radical prostatectomy and bilateral pelvic lymphadenectomy. The authors applied the multinomial log-linear regression analysis and found that 3 subgroups of PSA and 2 subgroups of biopsy Gleason sum allowed for patient stratification into 6 demarcated risk groups. The authors determined that the table could provide a basis for informed decision making and counseling of patients prior to surgery.

Regression models for ordinal responses take into consideration response variables with more than 2 categories and have some form of ordering. The work of Buch et al. (2005) on the effects of model choice on the relative risk estimates of blood cancer concluded with the continuation ratio model being best fit for the ordinal data in their study. The authors performed a case-control study to determine whether the estimated relative risks would vary when the multinomial, cumulative logit and continuation ratio models were applied to 4 response levels. These levels were non-blood cancer (controls), organ confined non-aggressive bladder cancer cases, organ confined aggressive bladder cancer cases and non-organ confined aggressive bladder cancer cases. Similar results were obtained by the multinomial logistic regression model compared to the multiple binary models. The cumulative logit model and the continuation ratio model gave different parameter estimate interpretations with the CR model being considered best fit based on the goodness-of-fit statistics, the regression diagnostic analysis, smaller standard errors and smaller 95% confidence intervals. Buch et al. (2005) stated that polychotomous logistic regression models are extensions of binary logistic models that preserve the natural order of outcomes with more than 2 categories, maximize the use of the available information in the categories, concurrently allow multiple logits to be estimated for nominal response variables, more statistical power is attained and estimators obtained are more efficient.

de Jong et al. (2019) presented a full-factorial simulation study to examine the predictive performance of the MLR models in relation to the relative size of the outcome categories, number of events and the number of events per variable. The authors carried out a study on ovarian cancer by applying penalized and unpenalized MLR to a development set (N = 2049) and a validation set (N = 2799) from a clinical study to produce a clinical prediction model. The model was to predict whether an ovarian tumor is benign, borderline malignant or invasive. de Jong et al. (2019) stated that due to the possibilities of invasive treatments worsening the prognosis of the ovarian cancer, it was vital to choose the most appropriate treatment using non-invasive procedures. The predictors of interest were age in years, presence of papillations with blood flow (yes/no), irregular cyst walls (yes/no), presence of acoustic shadows on the echo (yes / no), presence of ascites in the pouch of douglas (yes / no) and maximum diameter of solid component (continuous but no increase > 50mm). The authors utilized maximum likelihood, lasso and ridge to estimate the MLR models with the baseline category considered to be benign tumors. It was determined that in practice, the MLR model estimated by maximum likelihood tend to overfit even in samples with a high number of multinomial events per variable. Moreover, the authors discovered that the predictive performance of a MLR is affected by the sample size, analysis technique (whether or not shrinkage techniques are applied) and the relative size of the outcome variables.

According to de Jong et al. (2019), the calibration of the MLR models was at its worst when all outcome categories were equal in size, the multinomial events per variable was small and the number of predictors were low. However, when the multinomial events per variable was constant, model calibration improves as at least one of the outcome categories grows in size and as the number of predictors increases. The finding supports that as sample size affects the predictive performance , it affects the MLR model calibration. Additionally, the MLR estimated by ridge and lasso gave better calibration compared to the MLR estimated by maximum likelihood. Penalization reduces the overfit of estimates by inducing a small bias thus reducing the variance of the estimated probabilities. Improvement in calibration enhances the overall performance of the MLR model.de Jong et al. (2019) concluded that the predictive performance of a MLR model is very much related to the multinomial events per variable and the total sample size. When the total sample size is low, predictive performance can be poor and when the multinomial events per variable is below 10, the MLR model becomes at risk of serious miscalibration.

According to de Jong et al. (2019), the lasso MLR is applied to highly correlated predictors as performance increased significantly with increasing correlations. For the ridge MLR, correlation between predictors stabilizes the estimates, lowers the number of effective degrees of freedom as the coefficients are shrunk and improves the overall predictive performance.For the maximum likelihood MLR, when correlation changed, the ratio of events per effective degrees of freedom utilized remained constant which, in their study, may explain the small change noted in the model's predictive performance. The authors concluded that the MLR prediction model performance is related to multinomial events per variable and the total sample size and should be considered during development of the prediction model. Clearly, when the total sample size In addition, the work of Dos Santos and Berridge (2000) expounded on reasons a continuation ratio framework could ultimately be better than the multinomial and adjacent category logit models. The authors presented methods of analysis for ordinal repeated measures by investigating the influence of a group of explanatory variables on the overall condition of patients treated for breast cancer. The ordinal responses which were the patient's overall condition subsequent to surgery were death, progression of disease, disease static, partial response, complete response and no disease and the explanatory variables were effect of (clinical) stage, age and primary treatment (surgery). Briefly, the authors stated that the adjacent category logistic model and the baseline category models partition an ordinal response into 2 adjacent pairs with the adjacent category logistic model remaining invariant to changes in the grouping of the categories in the response variable. Thus, based on the continuation framework, the authors fitted a binary logistic, a binary logistic normal and a binary logistic non-parametric model and found that a deviance of $1803.59, 1800.77$ and 1794.73 respectively were attained by the 3 models. The binary logistic non-parametric model was found to be most appropriate as it accommodated the irreversibly of the ordinal response categories.

Conclusion

Following the literature, predictive modeling can enable predictive factors and prognostics to be utilized for accurate prescription of treatment to patients with chronic diseases. This study aims to utilize the available data to identify key predictive variables that will allow for the comparison of 3 regression models for ordinal responses for the prediction of the surgical stage of HIV infected and uninfected women with confirmed cervical cancer.

Chapter 3

Research Methodology

3.1 Introduction

The determination of the FIGO clinical stage of cervical cancer usually starts with a physical examination of the patient and is further supported by the availability of the pathological result of the collected cervical tissue. A review of the literature clearly shows that the clinical and laboratory parameters taken from women with confirmed diagnosis of cervical cancer are key to determining the surgical stage of women surgically treated for cervical cancer. This chapter aims to discuss the key predictive variables that can be utilized to predict the surgical stage of women with cervical cancer following surgical treatment and thus, develop, validate and compare the odds ratio of 3 regression models for ordinal responses. The parameters of each model were interpreted and the implication of each model evaluated on decision-making for the candidacy of a patient to undergo surgical treatment for cervical cancer. The details of the methods used are discussed and explained in this chapter.

3.2 Methods

Upon the study's approval, we adopted a cross-sectional design which utilized the retrospectively maintained database to identify all the women with International FIGO stage $0 - IVB$ cervical cancer managed by the Oncology department as outpatients at the MTRH's CCCDC from January 2014 to December 2018. Staging occurred according to the guidelines of the FIGO system; these did not change during the inclusion period.

The 3 regression models for ordinal responses that were developed, validated and compared were the Continuation Ratio model, the Polytomous (Multinomial) Logistic model and the Adjacent-Category Logistic model. The primary endpoint in this study was to build, validate and compare the odds ratio of 3 regression models for ordinal responses and establish which model best estimates and predicts the surgical stage after surgical treatment of cervical cancer. The model of choice should enable us to answer this question:

"Among the 3 regression models for ordinal responses, which model best estimates and predicts the surgical stage prior to surgical treatment of cervical cancer?"

"In order to gain a better insight into the possibility of predicting a patient's surgical stage prior to surgery based on the clinical stage and key explanatory variables, the R Studio statistical software was used to perform data analysis and construct the regression models. We carried out an inferential approach and a predictive approach. The inferential approach involved utilizing the original cervical cancer data of 75 women as the train data and simulating the validation (test) set of 10000 women with similar characteristics to the original data of 75 women. The predictive approach utilized the simulated data of 10000 women by partitioning 80% into the train set and 20% into the validation set. We were able to simulate the random sample of 10000 women with a weighted sampling scheme using the sample() function in R studio.

Descriptive statistics was carried out to get the holistic view of the characteristics of the women who sought treatment for cervical cancer at the Oncology department at the MTRH's Chandaria Cancer and Chronic Diseases Center.

Out of the 690 patients keyed into the cervical cancer database, only 75 were found to be eligible to be utilized in this study. The information from these patients was utilized to simulate a dataframe of 10000 patients with similar characteristics to the 75 patients. We were able to generate a random sample with a weighted sampling scheme by utilizing the sample() function in R studio.Proportions were estimated for each of the 75 women and provided for the weighted sampling. The response variable and the statistically significant predictors were considered during the simulation.We carried out an inferential and predictive approach which enabled the use of the original data and simulated data respectively. For the inferential approach, the original data was utilized as the train set in the development of the model and data simulated using the sample() function utilized as the validation set. For the predictive approach, data was simulated using the sample() function from the original data and the simulated data partitioned into the train and validation sets.

3.3 Regression models for Ordinal responses

Regression analysis investigates the influence of multiple predictors or independent variables on a dependent variable or outcome. In regression, an inferential approach draws conclusions from estimates and tests performed whereas a predictive approach tends to focus on past behaviors to enable better prediction of future outcomes.

The most commonly used model for the analysis of ordinal data is the Proportional or Cumulative Odds model which follows the assumption of proportional odds. The assumption of proportional odds in ordinal regression is that the effects of any explanatory variables are consistent or proportional across the different categories. Howbeit, Williams (2016) reports in his journal that the observed relationship between the predictors and the response variables are greatly misleading when the Cumulative Odds model assumptions is violated.

3.3.1 The Continuation Ratio Model

As a result of higher preference being given to the Cumulative Odds model for the analysis of ordinal data, Liu (2010) asserts that the Continuation Ratio model seems to be overlooked. Dolgun and Saracbasi (2014) states that this particular model is best suited in situations where the individual categories of the outcome variable are of intrinsic interest and are not arbitrary groupings of an underlying continuous variable. The CR model is applied whenever the focus of a study is on a particular category rather than at or below that category since a patient has to pass through a lower category before achieving a higher category. In this study, the ordinal categories for the CR model are successive stages of advancement in disease. According to Liu (2010), the CR model estimates the conditional probability $P(Y > j|Y \geq j)$ of being beyond a category $Y > j$ given that patient has achieved that particular category $Y = j$.
According to Allison, equation 3.1 shows the CR model as stated by Liu (2010), given a set of predictors:

$$P(Y > j|Y \geq j) = ln[\frac{P(Y \geq j|x_1, x_2, ..., x_p)}{P(Y = j|x_1, x_2, ..., x_p}] = \alpha_j + \beta_p X_p' \tag{3.1}$$

where $j = 1, 2, ..., j-1$ are the cut-points, $\beta_1, \beta_2, ..., \beta_p$ are the logit coefficients.

Also, the CR model estimates the odds of being in a particular category j relative to being in that category or beyond.
Equation 3.2 shows the CR model as stated by Liu (2010) where$P(Y = j|x_1, x_2, ..., x_p)$ is the conditional probability of being in category j:

$$ln[\frac{P(Y = j|x_1, x_2, ..., x_p)}{P(Y \geq j|x_1, x_2, ..., x_p}] = \alpha_j + (-\beta_p X_p') \tag{3.2}$$

where $j = 1, 2, ..., j-1$ are the cut-points, $\beta_1, \beta_2, ..., \beta_p$ are the logit coefficients. Consequently, the CR model estimates the probability of a response falling into a certain category (y_i) given that the person has been in that category or beyond.

3.3.2 Polytomous (Multinomial) Logistic Regression Model

The Pennsylvania State University expounds that binary logistic regression where $r = 2$ categories can be broadened to manage Y responses that are polytomous with $r > 2$ categories. This multi-category or polytomous responses can form a set of $r(r-1)/2$ logits or odds with only $r-1$ being non-redundant logits. The Polytomous logistic model is developed to predict the probabilities of the different possible outcomes of a categorical non-ordered response variable , given a set of predictor variables.

In a report by Starkweather and Moske (2011), the poytomous logistic regression model uses the maximum likelihood estimation to evaluate the probability of categorical membership. The author further states that the model's appeal is the non-assumption of normality, linearity and homoscedasticity, does not necessitate careful sample size consideration and examination of the outlying cases. Nonetheless, Starkweather and Moske (2011) insist that initial data analysis ought to be thorough with meticulous univariate, bivariate and multivariate evaluation. Moreover, multicollinearity among the independent variables should be evaluated and multivariate diagnostics utilized to assess the multivariate outliers and the exclusion of significant outliers. As stated by Starkweather and Moske (2011), the Polytomous Logistic regression model should satisfy these assumptions to give valid results:

1. Independence among the dependent response categories whereby membership in one category is not related to membership of another category.

2. Non-perfect separation. In order to avoid the estimation of unrealistic coefficients and great exaggeration of the effect size, the categories of the response variable are not perfectly separated by the predictors.

Grace-Martin (2019) describes the response variable in polytomous logistic regression as dummy coded into multiple 1 or 0 variables. If there are M categories, then there will be $M-1$ dummy variables meaning all but one category has its own dummy variable. She further expounds that each category's dummy variable has a value of 1 for its category and 0 for all the remaining categories. Notably, one category is considered as the reference category uniquely identified by all the other variables as 0 thus it does not need its own dummy variable. A separate binary regression model is estimated for each dummy variable resulting in $M-1$ binary regression models. The individual models each with its own intercept and regression coefficients will estimate the effect of the predictors on the probability of success in that specific category in comparison to the reference category.

According to Rodríguez (2019b), consider models where the probabilities Y_{ij}

depend on a vector x_i of covariates associated with the ith category. In addition, the maximum likelihood estimation of the parameters proceeds by maximization of the multinomial likelihood with the probabilities Y_{ij} viewed as functions of α_j and β_j parameters in equation 3.3.
Equation 3.3 shows the polytomous logit model as described by Rodríguez (2019b) to model the odds or logits:

$$\eta_{ij} = log\frac{Y_{ij}}{Y_{iJ}} = \alpha j + x_i'\beta_j \tag{3.3}$$

Similarly, Rodríguez (2019b) states that the Polytomous Logistic regression model can estimate the original probabilities by adopting the convention that $\eta_{ij} = 0$.
Equation 3.4 shows the polytomous logit model as described by Rodríguez (2019b) to model the odds or logits:

$$Y_{ij} = \frac{exp\{\eta_{ij}\}}{\Sigma_{k=1}^{J} exp\{\eta_{ij}\}} \tag{3.4}$$

where β_j is a vector of regression coefficients for $j = 1, 2, ..., j-1$ and α_j is a constant which brings about the assumption that the model matrix X does not include a column of ones.
However, some of the checks that need to be done prior to carrying out a multinomial logistic regression analysis as stated by Sharma (2019):

1. The response variable must be nominal. However, for multinomial regression, we need to run ordinal logistic regression.
2. The categorical predictor variables must be converted to dummy variables.
3. Multicollinearity must not be present.
4. A linear relationship must exist between the response variable and the continuous predictor variables.
5. Outliers and high influential points in the data must not be included.

Using R Studio, the mlogit package was utilized as it contains the functions to carry out the polytomous logistic regression analysis. The assumption of independence was tested by the Hausman-McFadden test. Starkweather and Moske (2011) explains that the t test determines if each coefficient is significantly different from zero and the Pseudo R_2 (McFadden R_2) is a measure of the size effect with the likelihood ratio test being the alternative test of goodness of fit.
In studies related to cervical cancer, polytomous models were utilized by Lee et al. (2015) to determine the odds of developing CIN 1 and CIN 2+ relative to women with negative Papanicolaou smear results. The authors developed and validated 4

cumulative risk score (CRS) schemes to improve Papanicolaou smear screening for women at high risk with high DNA load of high-risk HPV being the main predictor for CIN 1 and CIN 2 among other predictors. They determined that a simple CRS scheme that integrated HR-HPV testing and verified risk factors could effectively improve the identification of women at high risk of developing CIN2+. Lee et al. (2015) established that the greatest accuracy of cervical lesion was noted in the CRS schemes that included HPV testing.

3.3.3 The Adjacent-Category Logistic Model

In their article, Meisner et al. (2018) states that the set of logits generated by this model is analogous to the set of logits generated by the Polytomous model with the exception of the assuming a common β. The authors further report that this model takes advantage of the the ordinal outcome to attain parsimony and is more natural when interest lies in describing the effect of the predictor in terms of the odds in relation to particular outcome levels. Moreover, this model does not include cumulative probabilities thus an Adjacent-Category logistic model with separate effects (the potential crossing of the cumulative probability curves, which violates the ordering of the cumulative probabilities) is not a problem when utilizing this model.

Dolgun and Saracbasi (2014) states that this particular model is often utilized due to its close connection to log-linear models. Equation 3.5 as described by Ananth and Kleinbaum (1997) shows that the adjacent-category logistic model involves modeling the ratio of the two probabilities $P(Y = y_j)$ and $P(Y = y_{j+1}), j = 1, 2, ..., k$.

$$log[\frac{P(Y=y_j|x)}{P(Y=y_{j+1}|x)}] = \alpha_j - x'\beta_j, j = 1, 2, ..., k \tag{3.5}$$

Here, $\alpha_k = 0$ and $\beta_k = 0$. According to Ananth and Kleinbaum (1997), β_1 corresponds to the regression coefficient for the log-odds of $Y = y_1$ relative to $Y = y_2$, β_2 corresponds to the regression coefficient for the log-odds of $Y = y_2$ relative to $Y = y_3$, and so on with $(k - 1)$ intercept parameters α_j. Upon exponentiating the regression coefficient β_l for the l_{th} covariate x_l, we will obtain the odds ratio of comparing $Y = y_j$ versus $Y = y_{j+1}$ for a unit increase in x_l.

3.4 Procedure

The research design for this study was cross-sectional. The data for the study was retrospectively retrieved from the gynecological cervical cancer database. The data had been collected previously and was parallel to the patients' record files. The women who attend the gynecology clinic usually return for follow ups weekly, monthly and after 3 months. The gynecologists use files to record patient information at every visit and research assistants key in the recorded data into an MS access database at the close of the clinic sessions.

690 women with complete records sought treatment at the oncology clinic with only 75 women found to be eligible and their data utilized in the building of the predictive models. Moreover, data was simulated to test the performance of the developed models as the original data of 75 women was small to allow for partitioning. The independent variables in this study were age at first contact with the oncology team ranging from approximately $22 - 81$, parity of at least 2 live births per woman, international FIGO clinical stage which was dichotomized to clinical stages 1 and 2, HIV status of patient limited to either being HIV positive or HIV negative, vaginal involvement, parametrial involvement, marital status, weight of patient, smoking status, contraceptive use, method of cancer detection, biopsy pathology result, type of surgery done, symptoms which were either bleeding, vaginal discharge or lower abdominal pain, location of the cervical cancer tumour, grade of the tumour, the duration of the symptoms prior to diagnosis with the options being < 1 month, < 6 months, < 1 year, > 1 year and not stated with the dependent variable being surgical stage with the 3 categories being surgical stages $0, 1$ and 2.

Figure 3.1 show a flowchart displaying the surgical treatments that were availed and the surgical stage outcomes.

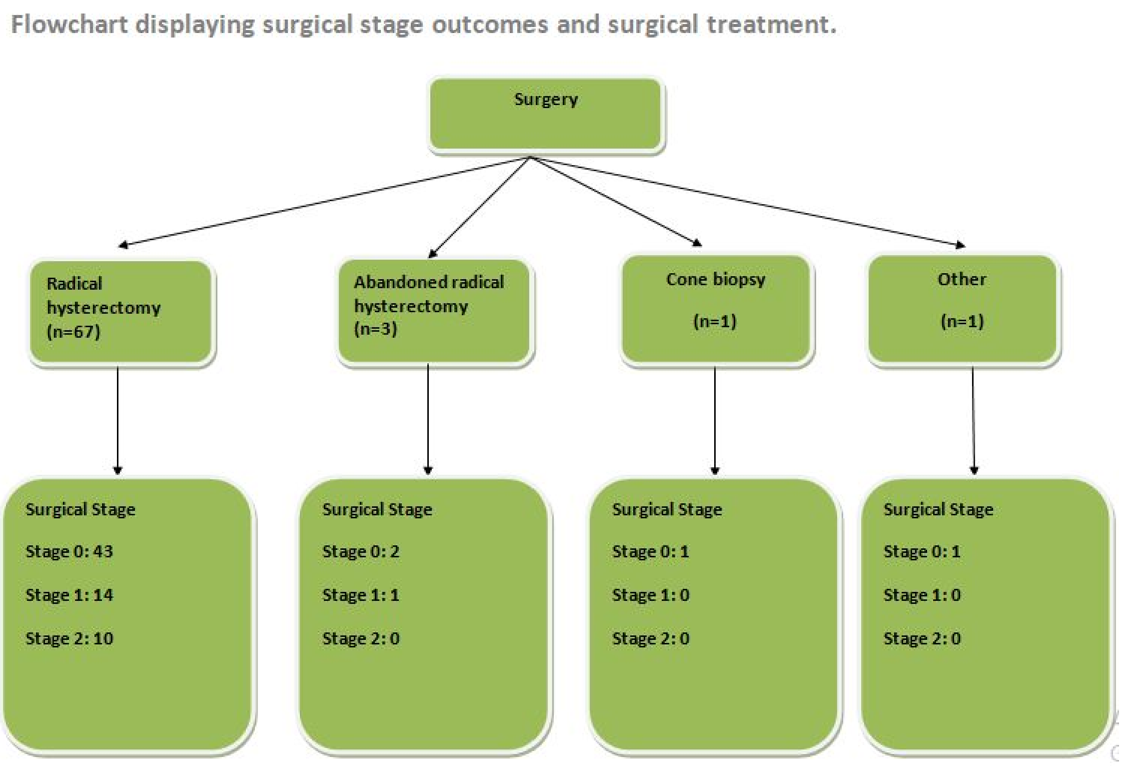

Figure 3.1: The surgical treatments availed and the surgical stage outcomes

Figure 3.2 is a flowchart displaying the surgical stage outcomes based on the colposcopic biopsy results.

Flowchart displaying cervical biopsy outcomes and FIGO clinical stages

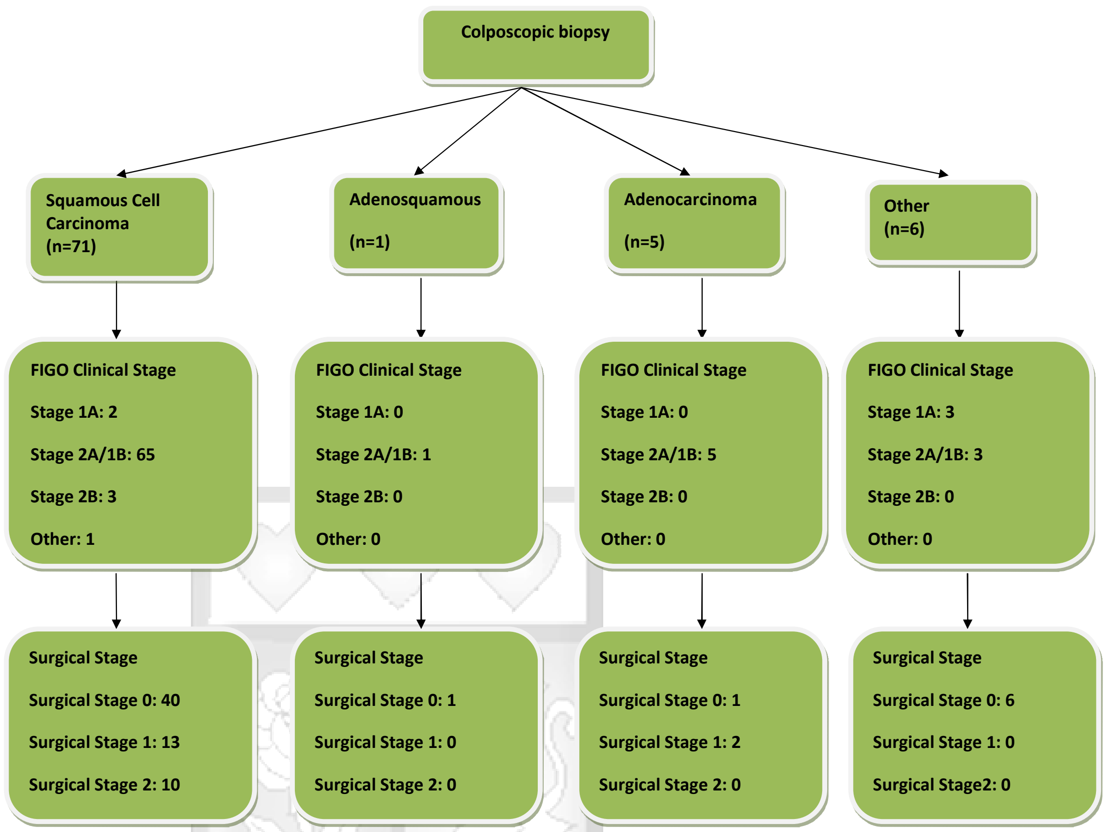

Figure 3.2: The colposcopic biopsy results and the surgical stage outcomes

3.4.1 Statistical tests

In this study, regression models were used to explore the relationship between the response variable (surgical stage) and the explanatory variables. The data was analyzed using R studio version 3.6.1. Chi-square tests and analysis of variance tests were carried out for categorical and numerical variables. The ANOVA test allowed us to examine the variation in the frequencies within each surgical stage (the response variable). Three regression models for ordinal data were developed and their predictive performance evaluated by comparing the odds ratios. These models were adapted because the response variable was an ordered variable. The 3 models were the multinomial (polytomous) logistic model, the continuation-ratio model and the adjacent-category logistic model for which the later 2 were developed with and without the proportional odds assumption.

We utilized R command multinom (Package: nnet) to fit 2 multinomial log-linear models via neural networks For the ACL model, we utilized the R vgam package that fits vector generalized and linear additive models appropriate to build the 2 adjacent-category models and the continuation ratio models both with and without proportional odds. We focused on the AIC goodness of fit statistic and the log likelihood ratios to compare the models. The response variable was coded as 0 for surgical stage 0, 1 for surgical stage 1 and 2 for surgical stage 2.

In their book, Meloun and Militky (2011) discussed the inferential tests for logistic regression being the tests of models and the tests of individual predictors. The author expounds that statistical inference is based on certain properties of maximum-likelihood estimators and on likelihood ratio tests. In our study, we have made comparisons between the null, univariate and multivariate models for the 3 regression models for ordinal responses. Among the several tests available, we utilized the likelihood ratio test and the deviance. The likelihood ratio test is the overall measure of how well the model fits and as stated by Meloun and Militky (2011), is similar to the residual or the error sum of squares value for multiple regression. The likelihood ratio test is the -2 times the log of the likelihood value (-2LL or -2 log likelihood) with a well-fitting model having a small value with the minimum value being 0 (-2LL $=0$). We can compare the likelihood value between equations whereby the difference represents the change in predictive fit. When the sample size is small or medium, Meloun and Militky (2011) reports that the likelihood ratio test is more accurate compared to the Wald test as shown in various simulation studies. Equation 3.6 defines the likelihood ratio test:

$$LR = -2[L_{subset} - L_{full}] = -2[ln(\frac{L_{subset}}{L_{full}})] \tag{3.6}$$

According to Meloun and Militky (2011), the likelihood ratio test is referred to as the Deviance when the full model in the likelihood ratio test is the saturated model (includes all possible terms including interactions). A saturated model allows for the predicted values from the model to be equal to the original data. In generalized linear models, the deviance statistic measures the deviance of the fitted GLM with respect to a saturated model for $\mathbb{E}[Y|X_1 = x_1, \ldots, X_p = x_p]$. A saturated model , also known as a perfect model is whereby the fitted responses ($\hat{Y}_i$) are the same as the observed responses (Y_i). Below is an example of a saturated model in logistic regression:

$$\hat{\mathbb{P}}[Y = 1|X_1 = X_{i1}, \ldots, X_k = X_{ip}] = Y_i, \quad i = 1, \ldots, n.$$

Additionally, the graphical representation 2.3 shows the saturated model and a fitted logistic regression model.

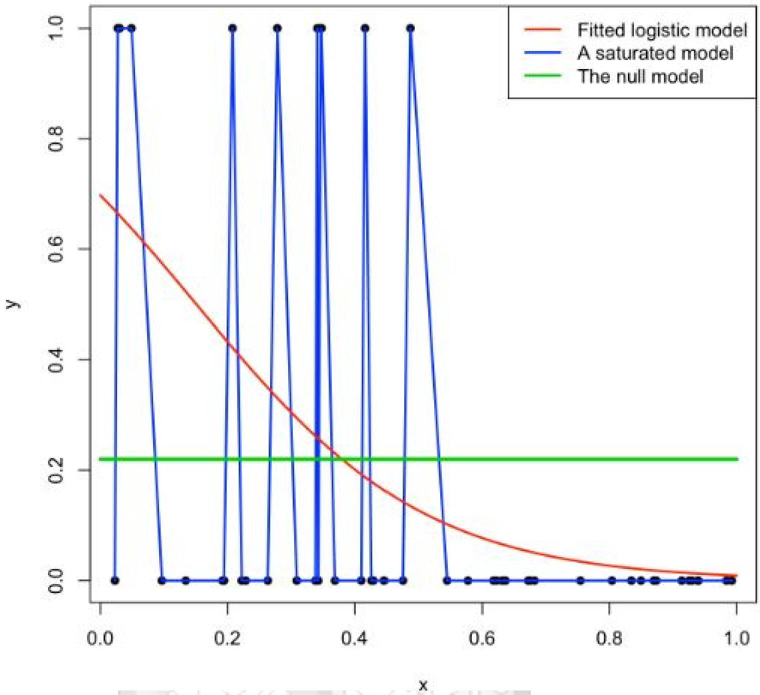

Figure 3.3: Graphical representation of logistic regression versus a saturated model and a null model

Portugués (2020) defines deviance as the difference of the log-likelihood between the fitted model $\ell(\hat{\boldsymbol{\beta}})$ and the saturated model, ℓ_s. When the μ_i is substituted by Y_i in equation ?? and $\theta_i = g(Y_i)$ if the canonical link function is utilized.
Equation ?? defines the log likelihood function and the deviance are shown as illustrated by Portugués (2020):

$$D = -2\left[\ell(\hat{\boldsymbol{\beta}}) - \ell_s\right]\phi.$$

The author further expounds that the log likelihood for the fitted model $\ell(\hat{\boldsymbol{\beta}})$ is always smaller than that of the saturated model ℓ_s. The deviance is equal to or larger than zero. A model that has a perfect fit will have a deviance of zero. Moreover, the deviance statistic is a generalization of the residual sum of squares of the linear model.

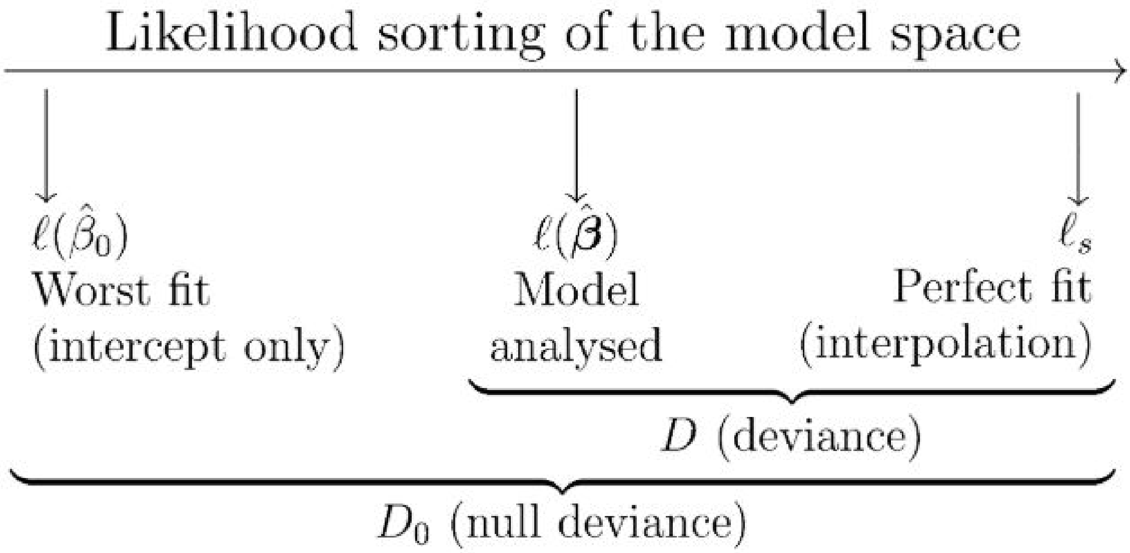

Figure 3.4: Graphical representation of logistic regression versus a saturated model and a null model

Figure 2.3 supports that the null deviance is the deviance of the worst model. The null model is considered the chosen benchmark for evaluating the scale of the deviance. The deviance and null deviance compare the models' improvement when predictors $X_1, \ldots, X_p$ are added and measure the percentage of explained deviance. Below is the equation for the null deviance as illustrated by Portugués (2020).

$$D_0 = \sum_{i=1}^{n} (Y_i - \hat{\eta}_i)^2 = \sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0\right)^2 = \text{SST},$$

In a similar manner, he states that the R^2 statistic is a measure or proportion of how well the model fits and a generalization of the determination coefficient for linear regression. $D = D_0$ and $R^2 = 0$ indicate that the predictors did not improve the regression and $D = 0$ and $R^2 = 1$ indicate a model that fits perfectly.
The data was analyzed using R studio version 3.6.1. Chi-square tests and analysis of variance tests were carried out for categorical and numerical variables. The ANOVA test allowed us to examine the variation in the frequencies within each surgical stage (the response variable). Three regression models for ordinal data were developed and their predictive performance evaluated by comparing the odds ratios. These models were adapted because the response variable was an ordered variable. The 3 models which were developed with and without the proportional odds assumption were the multinomial (polytomous)logistic model, the continuation-ratio model and the adjacent-category logistic model. We utilized R command multinom (Package: nnet) to fit 2 multinomial log-linear models via neural networks with and without

the proportional odds assumption. For the ACL model, we utilized the R vgam package that fits vector generalized and linear additive models appropriate to build the 2 adjacent-category models and the continuation ratio models both with and without proportional odds. We focused on the AIC goodness of fit statistic and the log likelihood ratios to compare the models. The response variable was coded as 0 for surgical stage 0, 1 for surgical stage 1 and 2 for surgical stage 2.

Chapter 4

Research Findings

4.1 Subjects of the study

In total, 87 women diagnosed with cervical cancer were found to fit the eligibility criteria between January 2014 and December 2018. We excluded women with unknown HIV status and incomplete follow up data within the stipulated period. Of these women, 61.4% and 18.0% were HIV infected and HIV uninfected women respectively. Based on the women who were on treatment to the end of the stipulated time period, a total of 311 and 312 women were clinically staged between stage $0 - 4B$ cervical cancer at baseline and at end of the time period respectively.

Table 1 below shows the characteristics of the women who underwent surgical treatment and whose records were found to be complete. At this particular time period,the characteristics of the women who were followed up from the start to the end of this period were analyzed.

Among the women who were surgically treated for cervical cancer, 87 women were surgically staged and considered eligible to participate in this study. The distribution of the surgical stage, which is the response variable was higher for surgical stage 0(66.67%) compared to surgical stage 1(20.00%) and surgical stage 2(13.33%). There was none classified under surgical stage 3 and 2 women classified under surgical stage 4 were excluded to minimize bias of the results. The mean womens' age and the median age at diagnosis was 46.61 and 46.00 and the frequently occurring clinical stage was 1B2 followed by 1B1. The predominant histologic type was squamous cell carcinoma at 63(86.3%). The patients diagnosed and listed as being with adenocarcinoma, adeno-squamous carcinoma and other were dropped from the analysis. The surgical treatments offered were radical hysterectomy, cone biopsy and abandoned radical hysterectomy with 1 unspecified. The type of surgical treatment was not statistically associated to the surgical stage outcomes. The mean overall weight was 64.76 kilograms (range: 0.00 – 163.00).

The marital status was classified as either single or married. The single patients comprised of the singles, widows, divorced and those who did not state their marital status. The majority of the women were married (76.8%) with the rest being single (23.2%) and the mean number of delivered pregnancies was approximately 5. Among the types of contraceptive methods being utilized, the depo provera method seems to be highly preferred by the women in this study. The international FIGO clinical stages were merged into clinical stages 1 and 2 and found to be 78.67% and 21.33% respectively with the clinical stage stated as 'others' being dropped. It became quite clear that on categorizing the FIGO clinical stages as 1 and 2 only, it was found to be statistically significant with a p-value of < 0.001. The major method of cervical cancer detection was based on patient symptoms with 44 (61.1%) women followed by 20 (27.8%) women who were diagnosed during cervical cancer screening testing . The method of detection was not statistically significantly associated with the surgical stage outcomes. Majority of the patients were non-smokers 73(97.3%) and was not statistically associated with the surgical stage outcomes. The location of the tumour was not statistically associated with the surgical stage outcomes with 48.5% of women having the tumour on both the endocervical and exocervical canal.

The predictors that were found to be statistically significant to the surgical stage were the FIGO clinical stage (p-value=$<$ 0.001), the vaginal involvement (p-value=$<$ 0.001), the parametrial involvement (p-value=0.008), symptomatic vaginal discharge (p-value=0.029) and symptomatic lower abdominal pain (p-value=0.048). Our results showed that the all other predictors were insignificant.

Table 1: Descriptive statistics for the predictor variables with a chi-square test and an anova test carried out for each categorical and numerical predictor respectively.

Surgical Stage	0 (N=50)	1 (N=15)	2 (N=10)	Total (N=75)	p-value
Parity					0.615
N-Miss	4	1	0	5	
Mean (SD)	4.652 (2.282)	5.357 (2.274)	4.800 (2.658)	4.814 (2.318)	
Median (Q1, Q3)	4.000 (3.000, 6.000)	5.500 (4.000, 6.000)	4.000 (3.000, 5.750)	4.000 (3.000, 6.000)	
Range	0.000 - 10.000	2.000 - 10.000	2.000 - 11.000	0.000 - 11.000	
Clinical Stage					< 0.001
1	43 (86.0%)	13 (86.7%)	3 (30.0%)	59 (78.7%)	
2	7 (14.0%)	2 (13.3%)	7 (70.0%)	16 (21.3%)	
Age at first clinical contact					0.595
N-Miss	3	0	0	3	
Mean (SD)	46.979 (11.709)	45.533 (8.790)	43.200 (9.762)	46.153 (10.858)	
Median (Q1, Q3)	46.000 (40.000, 53.500)	46.000 (43.000, 50.500)	41.500 (37.250, 49.000)	46.000 (40.000, 53.000)	
Range	22.000 - 81.000	24.000 - 59.000	27.000 - 59.000	22.000 - 81.000	
HIV Status					0.509
Negative	42 (84.0%)	13 (86.7%)	7 (70.0%)	62 (82.7%)	
Positive	8 (16.0%)	2 (13.3%)	3 (30.0%)	13 (17.3%)	
Vaginal Involvement					< 0.001
No	48 (96.0%)	14 (93.3%)	5 (50.0%)	67 (89.3%)	
Yes	2 (4.0%)	1 (6.7%)	5 (50.0%)	8 (10.7%)	
Parametrial Involvement					0.008
No	49 (98.0%)	13 (86.7%)	7 (70.0%)	69 (92.0%)	
Yes	1 (2.0%)	2 (13.3%)	3 (30.0%)	6 (8.0%)	

Marital Status					0.757
N-Miss	5	0	0	5	
Married	37 (82.2%)	11 (73.3%)	8 (80.0%)	56 (80.0%)	
Single	8 (17.8%)	4 (26.7%)	2 (20.0%)	14 (20.0%)	
Weight					0.690
Mean (SD)	66.200 (27.117)	63.867 (22.624)	58.900 (10.888)	64.760 (24.584)	
Median (Q1, Q3)	68.500 (56.500, 80.000)	67.000 (62.000, 75.500)	55.500 (51.750, 64.500)	67.000 (55.500, 77.500)	
Range	0.000 - 163.000	0.000 - 102.000	45.000 - 79.000	0.000 - 163.000	
Smoker					0.526
No	49 (98.0%)	14 (93.3%)	10 (100.0%)	73 (97.3%)	
Yes	1 (2.0%)	1 (6.7%)	0 (0.0%)	2 (2.7%)	
Contraception: None					0.725
No	36 (72.0%)	11 (73.3%)	6 (60.0%)	53 (70.7%)	
Yes	14 (28.0%)	4 (26.7%)	4 (40.0%)	22 (29.3%)	
Contraception: Condoms					
No	50 (100.0%)	15 (100.0%)	10 (100.0%)	75 (100.0%)	
Yes	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Contraception: Intrauterine.Device					0.274
No	45 (90.0%)	12 (80.0%)	10 (100.0%)	67 (89.3%)	
Yes	5 (10.0%)	3 (20.0%)	0 (0.0%)	8 (10.7%)	
Contraception: Oral Pill					0.122
No	46 (92.0%)	12 (80.0%)	7 (70.0%)	65 (86.7%)	
Yes	4 (8.0%)	3 (20.0%)	3 (30.0%)	10 (13.3%)	
Contraception: Depo Provera					0.838
No	37 (74.0%)	12 (80.0%)	7 (70.0%)	56 (74.7%)	
Yes	13 (26.0%)	3 (20.0%)	3 (30.0%)	19 (25.3%)	
					0.146

Method of Cancer Detection					
N-Miss	1	1	1	3	
Incidental	2 (4.1%)	1 (7.1%)	0 (0.0%)	3 (4.2%)	
Screening	18 (36.7%)	2 (14.3%)	0 (0.0%)	20 (27.8%)	
Symptoms	25 (51.0%)	10 (71.4%)	9 (100.0%)	44 (61.1%)	
Via	4 (8.2%)	1 (7.1%)	0 (0.0%)	5 (6.9%)	
Cervical biopsy Pathology result					0.245
N-Miss	2	0	0	2	
Adeno Carcinoma	1 (2.1%)	2 (13.3%)	0 (0.0%)	3 (4.1%)	
Adeno Squamous	1 (2.1%)	0 (0.0%)	0 (0.0%)	1 (1.4%)	
Other	6 (12.5%)	0 (0.0%)	0 (0.0%)	6 (8.2%)	
Squamous Cell	40 (83.3%)	13 (86.7%)	10 (100.0%)	63 (86.3%)	
Surgery Done					0.940
N-Miss	3	0	0	3	
Abandoned Radical Hysterectomy	2 (4.3%)	1 (6.7%)	0 (0.0%)	3 (4.2%)	
Cone Biopsy	1 (2.1%)	0 (0.0%)	0 (0.0%)	1 (1.4%)	
Other	1 (2.1%)	0 (0.0%)	0 (0.0%)	1 (1.4%)	
Radical Hysterectomy	43 (91.5%)	14 (93.3%)	10 (100.0%)	67 (93.1%)	
Symptom: Bleeding					0.169
No	21 (42.0%)	3 (20.0%)	2 (20.0%)	26 (34.7%)	
Yes	29 (58.0%)	12 (80.0%)	8 (80.0%)	49 (65.3%)	
Symptom: Discharge					0.029
No	34 (68.0%)	5 (33.3%)	4 (40.0%)	43 (57.3%)	
Yes	16 (32.0%)	10 (66.7%)	6 (60.0%)	32 (42.7%)	
Symptom: Pain					0.048
No	34 (68.0%)	5 (33.3%)	5 (50.0%)	44 (58.7%)	
Yes	16 (32.0%)	10 (66.7%)	5 (50.0%)	31 (41.3%)	
Tumour Location					0.347
N-Miss	31	6	5	42	
Both	7 (36.8%)	6 (66.7%)	3 (60.0%)	16 (48.5%)	
Endo-cervix	6 (31.6%)	0 (0.0%)	0 (0.0%)	6 (18.2%)	
Exo-cervix	5 (26.3%)	3 (33.3%)	2 (40.0%)	10 (30.3%)	
None	1 (5.3%)	0 (0.0%)	0 (0.0%)	1 (3.0%)	
Grade					0.576

N-Miss	2	0	0	2	
Grade 1	5 (10.4%)	0 (0.0%)	0 (0.0%)	5 (6.8%)	
Grade 2	12 (25.0%)	4 (26.7%)	2 (20.0%)	18 (24.7%)	
Grade 3	11 (22.9%)	4 (26.7%)	2 (20.0%)	17 (23.3%)	
Grade Not Stated	16 (33.3%)	7 (46.7%)	6 (60.0%)	29 (39.7%)	
N/A	4 (8.3%)	0 (0.0%)	0 (0.0%)	4 (5.5%)	
Symptoms Duration					0.322
N-Miss	23	3	2	28	
< 1 Months	2 (7.4%)	0 (0.0%)	0 (0.0%)	2 (4.3%)	
< 1 Year	3 (11.1%)	1 (8.3%)	3 (37.5%)	7 (14.9%)	
< 6 Months	8 (29.6%)	2 (16.7%)	3 (37.5%)	13 (27.7%)	
> 1 Year	1 (3.7%)	2 (16.7%)	1 (12.5%)	4 (8.5%)	
N/A	2 (7.4%)	0 (0.0%)	0 (0.0%)	2 (4.3%)	
Not Stated	11 (40.7%)	7 (58.3%)	1 (12.5%)	19 (40.4%)	

Table 4.1 lists the actual counts for the original (train) and simulation (test) data sets respectively utilized in the inferential analysis.

Category	Actual (train) dataset	Simulated (test) dataset
Surgical stage 0	50	6676
Surgical stage 1	15	2019
Surgical stage 2	10	1305

Table 4.1: Counts for patients in each surgical stage

4.2 Regression Analysis based on the Inferential approach

4.2.1 Polytomous (Multinomial) Logistic Regression Model

During the analysis, the reference category for the multinomial (polytomous) model was surgical stage 0. The 3 models fitted were the null model, univariate model and the multivariate model. We determined the presence of a relationship between the response variable and a combination of predictor variables by performing a likelihood ratio test. The results showed that the statistical significance of the probability of the model likelihood ratio (13.26) was 0.00132 indicating that there was a significant difference between the null (no predictors) and the univariate ML model (FIGO clinical stage only). Moreover, the probability of the model likelihood ratio (18.15) was 0.02013 showing a significant difference between the univariate and the multivariate ML models.

The ML model generalizes logistic regression to multiple outcome categories and thus, predicts the probabilities of the different possible outcomes of the outcomes given a set of predictor variables. The model attempts to explain the relative effect of the predictor variables on the surgical stage outcomes. Table 4.2 shows the null MLR model. The aim of the null model was to better understand the marginal distribution of the response variable in the absence of predictors.

Table 4.3 shows the fitted univariate MLR model with the international FIGO clinical stage as the sole predictor variable. The output shows that the log odds of being in surgical stage 1 compared to surgical stage 0 (reference category) decreased by 0.06 if moving from clinical stage 1 to clinical stage 2 and the log odds of being in surgical stage 2 compared to surgical stage 0(reference category) increased by 2.66 if moving from clinical stage 1 to clinical stage 2. The intercepts show the log odds for surgical stage 0 (reference category).

Table 4.4 shows the summary for a multivariate model built with the inclusion of the 5 statistically significant predictors. The first model compares surgical stage 1 to surgical stage 0 (reference category). We found that all the predictor variables were statistically insignificant and not associated to the surgical stage outcomes. The second model compares surgical stage 2 to surgical stage 0 where we found that only the FIGO clinical stage had a significant effect based on a p-value of 0.02297. The other 4 predictor variables were not significant to the surgical stage outcomes.

The log odds of being in surgical stage 1 compared to the surgical stage 0 will increase by 0.007 if moving from clinical stage 1 to clinical stage 2 and the log odds of being in surgical stage 2 compared to surgical stage 0 will increase by 2.401 if moving from clinical stage 1 to clinical stage 2. Thus, FIGO clinical stage exhibited positive regression coefficients and likelihood of increment with the higher categories of surgical stage. The log odds of being in surgical stage 1 compared to surgical stage 0 decreased by 0.359 if there was vaginal involvement observed during diagnosis and the log odds of being in surgical stage 2 compared to surgical stage 0 increased by 1.061 if there was vaginal involvement observed during diagnosis. The log odds of being in surgical stage 1 compared to surgical stage 0 increased by 1.509 and the log odds of being in surgical stage 2 compared to surgical stage 0 increased by 2.911 if the parametrium region was affected by the cervical cancer. The positive regression coefficients indicated that observed parametrial involvement was likely to lead to a higher category of surgical stage. The log odds of being in surgical stage 1 compared to surgical stage 0 increased by 1.261 and by 1.209 when a patient displayed symptomatic vaginal discharge and lower abdominal pain respectively. The log odds of being in surgical stage 2 compared to surgical stage 0 increased by 0.934 and decreased by 0.155 when a patient displayed symptomatic vaginal discharge and lower abdominal pain respectively.

	Coefficient	Standard Errors	z-statistic	p-value
Surgical Stage 1				
Intercept	-1.204	0.294	-4.10	0.00004
Surgical Stage 2				
Intercept	-1.610	0.350	-4.65	0.00000

Table 4.2: Summary for a null ML model.

	Coefficient	Standard Errors	z-statistic	p-value
Surgical Stage 1				
Intercept	-1.20	0.32	-3.78	0.00016
Clinical Stage 2	-0.06	0.86	-0.07	0.94757
Surgical Stage 2				
Intercept	-2.66	0.60	-4.46	0.00001
Clinical Stage 2	2.66	0.80	3.32	0.00089

Table 4.3: Summary for a univariate ML model built with the inclusion of the FIGO clinical stage predictor.

	Coefficient	Standard Errors	z-statistic	p-value
Surgical Stage 1				
Intercept	-1.20	0.32	-3.78	0.00016
Clinical Stage 2	-0.06	0.86	-0.07	0.94757
Vaginal Involvement: Yes	-0.359	1.623	-0.221	0.82509
Parametrial Involvement: Yes	1.509	1.354	1.114	0.26528
Symptoms Discharge: Yes	1.261	0.657	1.919	0.05498
Symptoms Pain: Yes	1.209	0.659	1.835	0.06651
Surgical Stage 2				
Intercept	-3.494	0.901	-3.878	0.00011
Clinical Stage 2	2.401	1.056	2.274	0.02297
Vaginal Involvement: Yes	1.061	1.183	0.897	0.36972
Parametrial Involvement: Yes	2.911	1.538	1.893	0.05836
Symptoms Discharge: Yes	0.934	0.910	1.026	0.30489
Symptoms Pain: Yes	-0.155	0.975	-0.159	0.89367

Table 4.4: Summary for a multivariate ML model built with the inclusion of the 5 statistically significant predictors.

Plot 4.1 features an effects plot which displays the predicted values of the surgical stage for given values of the predictor variables. The predicted value was obtained by inserting the predictor values into the ML multivariate model. The plot shows that there is a higher chances of being classified under surgical stage 0 for patients with FIGO clinical stage 1 and reduced chances for patients with FIGO clinical stage 2. In contrast, it was clear that the chances of being classified under surgical stage 2 were higher for patients with FIGO clinical stage 2 compared to patients with FIGO clinical stage 1. There was hardly any difference between the patients with FIGO clinical stage 1 and 2 when it came to being classified under surgical stage 1. Clearly, this supports the regression coefficient estimates from the univariate and multivariate ML models.

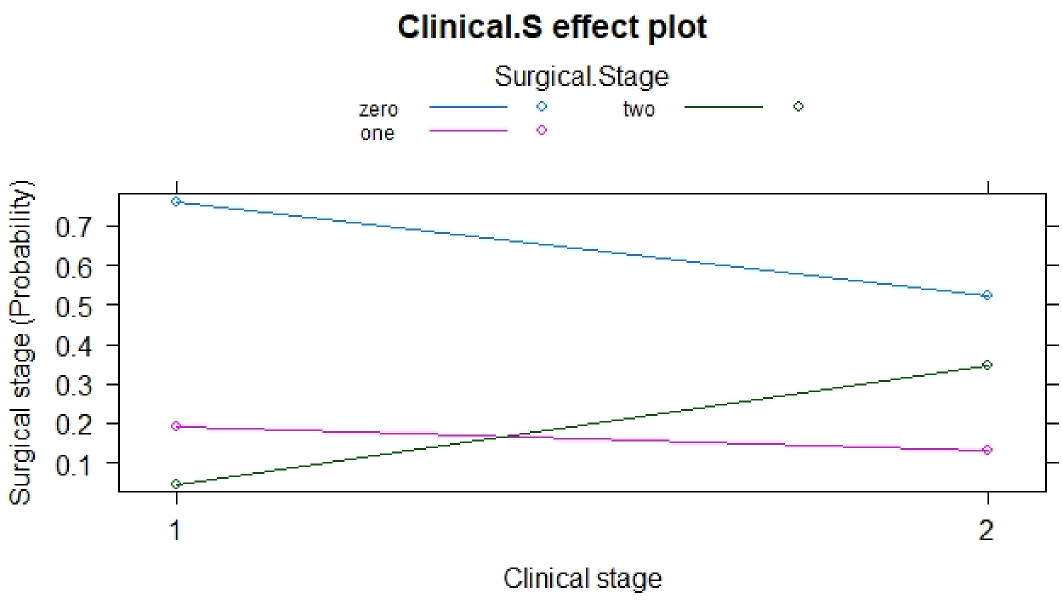

Figure 4.1: An Effects plot

Table 4.5 displays the goodness of fit statistics for the 3 models. The full ML model with the 5 statistically significant predictor variables had the lowest AIC and residual deviance of 121.72 and 97.72 respectively with the highest log likelihood ratio of -48.860. Thus, the multivariate multinomial logistic model was a better fit for the cervical cancer data compared to the univariate and null models.The presence of a relationship between the response variable and the combination of predictor variables was determined by a likelihood ratio test. The statistical significance of the probability of the model likelihood ratio (13.26) was 0.00132 shows that there was a significant difference between the null (no predictor variables) and the univariate model (FIGO clinical stage predictor only) . Also, the probability of the model likelihood ratio (18.15) was 0.02013 showing a significant difference between the univariate and multivariate models for the ML model.

	Deviance	Log Likelihood	AIC
Null ML model	115.87	-57.94	123.87
Univariate ML model	129.13	-64.56	133.13
Multivariate ML model	97.72	-48.86	121.72

Table 4.5: Goodness of fit statistics for the ML models

Equation 4.1 and 4.2 show the multivariate ML model for surgical stage 1 and surgical stage 2 both compared to the surgical stage 0 (the reference category). Based on the goodness-of-fit statistics, these 2 ML models were considered best in the classification of the surgical stage outcomes.

$$\begin{aligned} log[P(SS=1|SS=0)][= -1.200(0.320) - 0.060(0.860)clinicalstage- \\ 0.359(1.623)Vaginalinvolvement + 1.509(1.354) \\ Parametrialinvolvement + 1.261(0.657) \\ Symptom: Discharge + 1.209(0.659)Symptom: Pain] \end{aligned} \tag{4.1}$$

$$\begin{aligned} log[P(SS=2|SS=0][= -3.494(0.901) + 2.401(1.056)clinicalstage \\ + 1.061(1.183)Vaginalinvolvement + 2.911(1.538)Parametrialinvolvement \\ + 0.934(0.910)Symptom: Discharge - 0.155(0.975)Symptom: Pain] \end{aligned} \tag{4.2}$$

Confusion matrices 4.6 and 4.7 derived for the train and validation set gave an accuracy of $70\%[95\%CI : 59.02\% - 80.62\%]$ and $62.97\%[95\%CI : 62.01\% - 63.92\%]$. With reference to table 4.8, the sensitivity for surgical stage 0, 1 and 2 was $0.960, 0.0667$ and 0.4000 for the train set and $0.926, 0.0521$ and for 0.0223 for the validation set. The specificity for surgical stage 0, 1 and 2 was $0.280, 0.983$ and 0.954 for the train set and $0.073, 0.974$ and for 0.953 for the validation set.

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	48	0	2
Surgical stage one	13	1	1
Surgical stage two	5	1	4

Table 4.6: Confusion matrix of the multivariate ML model based on the train set.

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	45	39	172
Surgical stage one	90	68	320
Surgical stage two	1884	1198	6184

Table 4.7: Confusion matrix for the multivariate ML model based on the validation set

	Sensitivity	Specificity	Prevalance
Train set			
Surgical stage zero	0.9600	0.2800	0.6667
Surgical stage one	0.0667	0.9833	0.2000
Surgical stage two	0.4000	0.9539	0.1333
Validation set			
Surgical stage zero	0.9263	0.0728	0.6676
Surgical stage one	0.0223	0.9736	0.2019
Surgical stage two	0.0521	0.9529	0.1305

Table 4.8: Sensitivity, specificity and Prevalance for the multivariate ML model based on the train and validation set respectively.

With the best fit model being the multivariate ML model, the odds ratios were extracted from the models. The odds ratios 4.9 extracted from the multivariate multinomial logistic model which displayed the best fit model for the cervical cancer data. The odds of being classified into surgical stage 1 over surgical stage 0 was 1.01 [CI: 0.13 – 7.61] higher for patients diagnosed with FIGO clinical stage 2 versus those diagnosed with FIGO clinical stage 1 while holding all other predictors constant. Moreover, the odds of being classified into surgical stage 2 over surgical stage 0 was 11.03 CI:[1.39 – 87.36] times higher for patients diagnosed with FIGO clinical stage 2 versus those diagnosed with FIGO clinical stage 1 while holding all other predictors constant. The odds of being classified into surgical stage 1 over surgical stage 0 was 0.70 CI:[0.13 – 7.61] times lower and in contrast,the odds of being in surgical stage 2 over surgical stage 0 was2.89 CI: [0.28 – 29.35] times higher for patients with the vaginal region observed to be affected by the cancer during diagnosis versus those without any vaginal involvement while holding all other predictors constant. The odds of being classified into surgical stage 1 over surgical stage 0 was 4.52 CI: [0.3 – 64.28] times higher and the odds of being classified into surgical stage 2 over surgical stage 0 was 18.38 CI: [0.9 – 374.72] times higher for patients with the parametrial region affected by the cancer versus those without any parametrial involvement while holding other predictors constant. The odds of being classified into surgical stage 1 over surgical stage 0 was 3.53 CI: [0.97 – 12.78] times higher and the odds of being classified into surgical stage 2 over surgical stage 0 was 2.54 CI: [0.43 – 15.13] times higher for the patients with symptomatic vaginal discharge during diagnosis versus those without the symptomatic vaginal discharge, holding all other predictors constant. The odds of being classified into surgical stage 1 over surgical stage 0 was 3.35 CI:[0.92 – 12.18] times higher and in contrast, the odds of being into surgical stage 2 over surgical stage 0 was 0.86 CI:[0.13 – 5.79] times lower for the patients with symptomatic lower abdominal pain versus those without any pain, holding all other predictors constant.

	Intercept	Clinical Stage 2	Vaginal Involvement	Parametrial Involvement	Symptoms Discharge	Symptoms Pain
Surgical Stage 1						
Coefficient	-2.499	0.007	-0.359	1.509	1.261	1.209
Std Error	0.622	1.032	1.623	1.354	0.657	0.659
Z statistic	-4.018	0.007	-0.221	1.114	1.919	1.835
p-value	<0.01	0.9944	0.82509	0.26528	0.05498	0.06651
OR (95% CI)	0.08(0.02, 0.28)	1.01(0.13, 7.61)	0.07(0.03, 16.83)	4.52(0.32, 64.28)	3.53(0.97, 12.78)	3.35 (0.92, 12.18)
Surgical Stage 2						
Coefficient	-3.494	2.401	1.061	2.911	0.934	-0.155
Std Error	0.901	1.056	1.183	1.538	0.91	0.975
Z statistic	-3.878	2.274	0.897	1.893	1.026	-0.159
p-value	<0.001	0.02297	0.36972	0.05836	0.30489	0.87367
OR (95% CI)	0.03(0.01,0.18)	11.03(1.39,87.36)	2.89(0.28,29.35)	18.38(0.9,374.72)	2.54(0.43,15.13)	0.86(0.13, 5.79)

Table 4.9: The table of the odds ratios extracted from the multivariate ML model which displayed the best fit model for the cervical cancer data.

4.2.2 Continuation Ratio Model

When the focus is on a particular category given that a patient must pass through a lower surgical stage category before achieving a higher category, the continuation ratio model is considered a more appropriate choice. The proportional odds assumption was tested by fitting this particular model with and without the proportional odds assumption.

The tables below show the output for the CR null, univariate and multivariate models with and without proportional odds respectively. The univariate CR model with and without proportional odds output considered the FIGO clinical stage as the only predictor. Clearly, the CR univariate model with proportional odds showed that the FIGO clinical stage had a significant effect on the surgical stage response with a p-value of 0.000892 when the regression coefficient was common. The estimated log regression coefficient for the FIGO clinical stage , $\beta = 1.649(0.496), z = 3.322$ and $p < 0.05$ showed that the FIGO clinical stage upon diagnosis had a positive effect on the surgical stage responses. The CR model without proportional odds gave separate effects. The FIGO clinical stage predictor variable was found to be statistically significant. For having a surgical stage outcome greater than surgical stage 1 relative to being in surgical stage 1, the estimated log regression coefficient for the FIGO clinical stage was $\beta = 1.240(0.583)$, z-value= 2.127 and a p-value of 0.0334 indicating that the FIGO clinical stage had a significant positive effect on the surgical stage 1 responses. In addition, having a surgical stage outcome greater than surgical stage 2 relative to being in surgical stage 2, the estimated logit regression coefficient for FIGO clinical stage was $\beta = 2.719(1.026)$, z-value=2.650 and a p-value of 0.00805 indicating that the FIGO clinical stage had a significant positive effect on surgical stage 2 responses.

The multivariate CR model with and without proportional odds output with the 5 predictors that were found to be statistically significant were included.For the CR multivariate model with proportional odds, only the FIGO clinical stage estimated log regression coefficient, $\beta = 1.449(0.632)$, z-value=2.293 had a positive effect on the surgical stage responses. In addition, a p-value of 0.02182 showed that the FIGO clinical stage is a statistically significant predictor for having a surgical stage outcome greater than surgical stage 1 relative to being in surgical stage 1 as well as having a surgical stage outcome greater than surgical stage 2 relative to being in surgical stage 2. The remaining 4 predictors that were not statistically significant to the surgical stage responses were vaginal involvement, parametrial involvement, symptomatic vaginal discharge and lower abdominal pain. For the CR multivariate model without proportional odds, we got separate effects for the surgical stage responses.For having a surgical stage outcome greater than surgical stage 1 relative to being in surgical stage 1, the symptomatic vaginal discharge was found to have a positive effect and was significant with an estimated logit coefficient $\beta = 1.103(0.568)$, z-value=1.943 and a p-value of 0.052. The estimated logit coefficients for FIGO clinical stage, vaginal involvement,parametrial involvement and symptomatic lower abdominal pain were not statistically significant and had no effect . Also, it was clear that for having a surgical stage outcome greater than surgical stage 2 relative to being in surgical stage 2, the FIGO clinical stage had a positive effect and was statistically significant with an estimated logit regression coefficient of $\beta = 3.833(1.817)$, z-value=2.109 and p-value of 0.0349. The estimated logit regression coefficients for vaginal involvement,parametrial involvement, symptomatic vaginal discharge and symptomatic lower abdominal pain were not statistically significant and had no effect.

	Coefficient	Standard Errors	z-statistic	p-value
CR Model with Proportional Odds				
Intercept1	-0.693	0.245	-2.830	< 0.00466
Intercept2	-0.405	0.40892	-0.993	< 0.32062
CR Model without Proportional Odds				
Intercept1	-0.693	0.245	-2.830	< 0.00466
Intercept2	-0.405	0.408	-0.993	0.32062

Table 4.10: The Summary for a CR null model.

	Coefficient	Standard Errors	z-statistic	p-value
CR Model with Proportional Odds				
Intercept1	-1.092	0.288	-3.790	< 0.00015
Intercept2	-1.042	0.492	-2.115	< 0.03443
Clinical Stage 2	1.649	0.496	3.322	0.00089
CR Model without Proportional Odds				
Intercept1	-0.989	0.293	-3.376	< 0.00074
Intercept2	-1.466	0.641	-2.289	0.02206
Clinical Stage 1	1.240	0.583	2.127	0.03339
Clinical Stage 2	2.719	1.026	2.650	0.00806

Table 4.11: The Summary for a CR univariate model with the inclusion of the FIGO clinical stage predictor.

	Coefficient	Standard Errors	z-statistic	p-value
CR Model with Proportional Odds				
Intercept1	-1.773	0.441	-4.021	< 0.01
Intercept2	-2.254	0.698	-3.229	< 0.01
Clinical Stage 2	1.449	0.632	2.293	0.022
Vaginal Involvement: Yes	0.982	0.868	1.131	0.258
Parametrial Involvement: Yes	1.61	0.882	1.825	0.068
Symptoms Discharge: Yes	0.717	0.498	1.438	0.1504
Symptoms Pain: Yes	0.382	0.507	0.754	0.4508
CR Model without Proportional Odds				
Intercept1	-2.05	0.512	-4.001	< 0.01
Intercept2	-1.195	1.239	-0.965	0.335
Clinical Stage 1	1.001	0.787	1.271	0.204
Clinical Stage 2	3.833	1.817	2.109	0.035
Vaginal Involvement: Yes1	0.736	1.161	0.634	0.526
Vaginal Involvement: Yes2	1.584	1.699	0.933	0.351
Parametrial Involvement: Yes1	1.829	1.223	1.495	0.131
Parametrial Involvement: Yes2	3.220	1.918	1.679	0.093
Symptoms Discharge: Yes1	1.103	0.568	1.943	0.052
Symptoms Discharge: Yes2	-1.412	1.542	-0.916	0.360
Symptoms Pain: Yes1	0.823	0.577	1.427	0.154
Symptoms Pain: Yes2	-0.832	1.515	-0.209	0.227

Table 4.12: The Summary for a CR multivariate model with the inclusion of the 5 predictors for Surgical Stage 1 and Stage 2.

Table 4.13 displays the goodness of fit statistics for the 3 models. An AIC of 118.899 shows that the multivariate CR model without proportional odds gave the best fit for the cervical cancer data with further confirmation based on a residual deviance and log likelihood ratio of 94.89 and -47.44 respectively. Moreover, the VGAM likelihood ratio test was carried out for the 2 CR multivariate models and a chi-square p-value of 0.08023 showed that the fit was not significantly different and thus, the multivariate CR model without proportional odds was found to be adequate.

	Deviance	Log Likelihood	AIC
Null CR model	129.13	-64.56	133.13
Univarite CR model	117.56	-58.78	123.56
Multivariate CR model	104.72	-52.36	118.72
	Deviance	Log Likelihood	AIC
Null CR model	129.13	-64.56	133.13
Univariate CR model	115.87	-57.94	123.87
Multivariate CR model	94.89	-47.44	118.89

Table 4.13: Goodness of fit statistics for the CR models with and without Proportional Odds respectively

Equation 4.3 and 4.4 show the multivariate CR model without proportional odds assumptions for surgical stage 1 and surgical stage 2. Based on the goodness-of-fit statistics, these 2 CR models without proportional odds were considered best in the classification of the surgical stage outcomes.

$$\begin{aligned} log[P(SS = 1|SS \geq 1)][= &-2.050(0.512) + 1.001(0.787)clinicalstage+ \\ &0.736(1.161)Vaginalinvolvement + 1.829(1.223) \\ &Parametrialinvolvement + 1.103(0.568) \\ &Symptom : Discharge + 0.823(0.577)Symptom : Pain] \end{aligned} \tag{4.3}$$

$$\begin{aligned} &log[P(SS = 2|SS \geq 2][= -1.1950(1.239) + 3.833(1.817)clinicalstage \\ &+ 1.584(1.699)Vaginalinvolvement + 3.220(1.918)Parametrialinvolvement \\ &- 1.412(1.542)Symptom : Discharge - 1.832(1.515)Symptom : Pain] \end{aligned} \tag{4.4}$$

The 3 CR models with and without proportional odds were compared to determine the model best fit for the cervical cancer data. The fitted multivariate CR model with proportional odds had a misclassification rate of 29.33% and 37.74% whereas the fitted multivariate CR model without proportional odds had a misclassification rate of 30.67% and 39.09% when the train and validation data sets were utilized respectively.

Confusion matrices 4.14 and 4.15 derived for the train and validation set from the multivariate CR model without proportional odds gave an accuracy of $69.33\%[95\%CI : 57.62\% - 79.47\%]$ and $60.91\%[95\%CI : 59.95\% - 61.87\%]$. With reference to table 4.16, the sensitivity for surgical stage 0, 1 and 2 was 0.734, 0.333 and 0.500 for the train set and 0.668, 0.227 and 0.132 for the validation set. The specificity for surgical stage 0, 1 and 2 was $0.727, 0.806$ and 0.910 for the train set and $0.338, 0.799$ and for 0.870 for the validation set.

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	47	1	2
Surgical stage one	12	1	2
Surgical stage two	5	1	4

Table 4.14: Confusion matrix of the multivariate CR model without proportional odds based on the train set.

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	5883	329	464
Surgical stage one	1779	119	121
Surgical stage two	1140	76	89

Table 4.15: Confusion matrix for the multivariate CR model without proportional odds based on the validation set

	Sensitivity	Specificity	Prevalance
Train set			
Surgical stage zero	0.9600	0.2800	0.6667
Surgical stage one	0.0667	0.9833	0.2000
Surgical stage two	0.4000	0.9539	0.1333
Validation set			
Surgical stage zero	0.6684	0.3381	0.8802
Surgical stage one	0.2271	0.7995	0.0524
Surgical stage two	0.1320	0.8696	0.0674

Table 4.16: Sensitivity, Specificity and Prevalance for the multivariate CR model without proportional odds based on the train and validation set respectively.

Table 4.17 shows the extracted odds ratios for the best fit model which was found to be the CR model without proportional odds.

Predictor variables	Odds ratio	CI: 2.5 %	CI: 97.5 %
(Intercept):1	0.13	0.05	0.35
(Intercept):2	0.30	0.03	3.43
Clinical.Stage 2:1	2.72	0.58	12.72
Clinical.Stage 2:2	46.20	1.31	1627.54
Vaginal.Involvement Yes:1	2.09	0.21	20.30
Vaginal.Involvement Yes:2	4.88	0.17	136.11
Parametrial.Involvement Yes:1	6.23	0.57	68.45
Parametrial.Involvement Yes:2	25.02	0.58	1073.25
Symptoms...Discharge Yes:1	3.01	0.99	9.16
Symptoms...Discharge Yes:2	0.24	0.01	5.00
Symptoms...Pain Yes:1	2.28	0.74	7.06
Symptoms...Pain Yes:2	0.16	0.01	3.12

Table 4.17: Odds Ratios for the CR model without Proportional Odds

For the model 4.3 $logit[P(SS = 1|SS \geq 1]$ and model 4.4 $logit[P(SS = 2|SS \geq 2]$ respectively, the odds of having an outcome greater than surgical stage 1 relative to being in surgical stage 1 was $2.72[0.58 - 12.72]$ times higher and the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $46.20[1.31 - 1627.54]$ times higher among the patients diagnosed with FIGO clinical stage 2 compared to the patients diagnosed with FIGO clinical stage 1, after controlling for the effects of other predictors in the model. In addition, the odds of having an outcome greater than surgical stage 1 relative to being in surgical stage 1 was $2.09[0.21 - 20.30]$ times higher and the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $4.88[0.17 - 136.11]$ times higher among the patients considered to have the vaginal region affected by the cancer (vaginal involvement) compared to the patients without any vaginal involvement after controlling for the effects of other predictors in the model. The odds of having an outcome greater than surgical stage 1 relative to being in surgical stage 1 was $6.23[0.57 - 68.45]$ times higher and the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $25.02[0.58 - 1073.25]$ times higher among the patients considered to have the parametrium region affected by the cervical cancer (parametrial involvement) compared to the patients without any parametrial involvement after controlling for the effects of other predictors in the model. The odds of having an outcome greater that surgical stage 1 relative to being in surgical stage 1 was $3.01[0.99 - 9.16]$ times higher and in contrast, the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $0.24[0.01 - 5.00]$ times lower among the patients with symptomatic vaginal discharge (Symptoms: Discharge) compared to the patients who did not have symptomatic vaginal discharge after controlling for the effects of other predictors in the model. The odds of having an outcome greater than surgical stage 1 relative to being in surgical stage 1 was $2.28[0.74 - 7.06]$

times higher and in contrast, the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $0.16[0.01 - 3.12]$ times lower among the patients displaying symptomatic lower abdominal pain (Symptoms: Pain) compared to the patients without symptomatic lower abdominal pain after controlling for the effects of other predictors in the model.

4.2.3 Adjacent Category Logistic Regression Model

The Adjacent Category Logit model Ripley et al. (2016) is a special form of generalized logit models that involves the simultaneous estimation of the effects of predictor variables in pairs of adjacent categories. The ACL model involves the ratio of two probabilities $P[Y = y_j]$ and $P[Y = y_{j+1}]$. The proportional odds assumption was tested by fitting the ACL model with and without the proportional odds assumption.

The tables below shows the summary of the ACL null, univariate and multivariate models with and without proportional odds. For the ACL univariate model with proportional odds, we found that the FIGO clinical stage had a statistically significant effect on the surgical stage response with a p-value of 0.00207. The estimated logit regression coefficient for the FIGO clinical stage, $\beta = -1.1740$, z-value=-3.080 and a p-value < 0.05 which showed that the FIGO clinical stage upon diagnosis had a negative effect on each adjacent surgical stage response category (common regression coefficients). The ACL univariate model without proportional odds gave separate effects. We found that the estimated logit regression coefficient $\beta = -2.719$, z-value=-2.650 and the p-value=0.008057 indicated that the log odds of being in surgical stage 2 versus surgical stage 1 was -2.719 when the FIGO clinical stage increased by 1 unit, holding all other predictors constant. Thus, the FIGO clinical stage had a significant effect on the probability of being in surgical stage 2 versus surgical stage 1. However, the FIGO clinical stage had no significant effect on the probability of being in surgical stage 1 versus surgical stage 0 with an estimated logit regression coefficient $\beta = 0.057$, z-value=0.066 and the p-value=0.9477.

For the ACL multivariate model with proportional odds , only the FIGO clinical stage with an estimated logit regression coefficient $\beta = -1.044(0.509)$ and a z-value=-2.05 had a negative effect on the surgical stage responses. Moreover, a p-value of 0.04036 confirmed that it is a statistically significant predictor. The remaining 4 predictors that were not statistically significant to the surgical stage responses were vaginal involvement, parametrial involvement, symptomatic vaginal discharge and lower abdominal pain. For the ACL multivariate model without proportional odds, we get separate effects for the surgical stage responses. The FIGO clinical stage had a negative effect on the probability of being classified under surgical stage 2 versus surgical stage 1. The estimated logit regression coefficient $\beta = -2.349(1.258)$, z-value $= -1.903$ and a p-value of 0.057 indicates that it is not a significant predictor and the log-odds of being classified under surgical stage 2 versus surgical stage 1 was -2.394 when the FIGO clinical stage increased by 1 unit, holding all other predictors constant. In addition, the symptomatic vaginal discharge predictor had a negative effect on the on the probability of being classified under surgical stage 1 versus surgical stage 0. As indicated by the estimated logit regression coefficient $\beta = -1.261(0.657)$ and z-value$=-1.920$, the log odds of being classified under surgical stage 1 versus surgical stage 0was -1.261. A p-value of 0.05488 shows that the predictor was not significant to the surgical stage response. Clearly, of the 5 predictor variables, we found no statistically significance to the probability of being classified either under surgical stage 1 versus surgical stage 0 or surgical stage 2 versus surgical stage 1.

	Coefficient	Standard Errors	z-statistic	p-value
ACL Model with Proportional Odds				
Intercept1	1.204	0.294	4.090	< 0.0000
Intercept2	0.405	0.408	0.993	< 0.3206
ACL Model without Proportional Odds				
Intercept1	1.204	0.294	4.090	< 0.0000
Intercept2	0.405	0.408	0.993	0.3206

Table 4.18: The Summary for a ACL null model .

	Coefficient	Standard Errors	z-statistic	p-value
ACL Model with Proportional Odds				
Intercept1	1.434	0.306	4.688	< 0.0000
Intercept2	0.914	0.485	1.887	< 0.0592
Clinical Stage 2	-1.174	0.381	-3.080	0.0021
ACL Model without Proportional Odds				
Intercept1	1.196	0.317	3.779	< 0.0002
Intercept2	1.466	0.640	2.289	0.0221
Clinical Stage 1	0.057	0.862	0.066	0.9478
Clinical Stage 2	-2.719	1.026	-2.650	0.0081

Table 4.19: The Summary for a ACL univariate model with the inclusion of the FIGO clinical stage predictor.

	Coefficient	Standard Errors	z-statistic	p-value
ACL Model with Proportional Odds				
Intercept1	1.937	0.397	4.874	< 0.0000
Intercept2	1.861	0.674	2.760	< 0.0058
Clinical Stage 2	-1.044	0.509	-2.050	0.0404
Vaginal Involvement: Yes	-0.624	0.661	-0.943	0.3456
Parametrial Involvement: Yes	-1.192	0.676	-1.762	0.078
Symptoms Discharge: Yes	-0.668	0.409	-1.635	0.1021
Symptoms Pain: Yes	-0.330	0.410	-0.806	0.4200
ACL Model without Proportional Odds				
Intercept1	2.499	0.622	4.017	< 0.0001
Intercept2	0.994	1.044	0.953	0.3407
Clinical Stage 1	-0.006	1.032	-0.006	0.9950
Clinical Stage 2	-2.394	1.258	-1.903	0.057
Vaginal Involvement: Yes1	0.358	1.623	0.221	0.8253
Vaginal Involvement: Yes2	-1.420	1.552	-0.915	0.360
Parametrial Involvement: Yes1	-1.509	1.354	-1.114	0.2652
Parametrial Involvement: Yes2	-1.403	1.444	-0.972	0.3313
Symptoms Discharge: Yes1	-1.261	0.657	-1.920	0.0549
Symptoms Discharge: Yes2	0.327	1.027	0.319	0.7501
Symptoms Pain: Yes1	-1.209	0.659	-1.835	0.067
Symptoms Pain: Yes2	1.364	1.063	1.283	0.1996

Table 4.20: The Summary for a ACL multivariate model with the inclusion of the 5 predictors for Surgical Stage 1 and Stage 2.

Table 4.21 displays the goodness of fit statistics for the 3 models. An AIC of 121.72 indicates that the multivariate ACL model without proportional odds gave the best fit for the cervical cancer data with further confirmation based on a residual deviance and log likelihood ratio of 97.72 and -48.86 respectively. We carried out a VGAM likelihood ratio test for the 2 multivariate ACL models and a chi-square p-value of 0.002981 indicating that both fits were significantly different from each other.

	Deviance	Log Likelihood	AIC
Null ACL model	129.13	-64.56	133.13
Univariate ACL model	119.03	-59.51	125.03
Multivariate ACL model	106.13	-53.07	120.13
	Deviance	Log Likelihood	AIC
Null ACL model	129.13	-64.56	133.13
Univariate ACL model	115.87	-57.94	123.87
Multivariate ACL model	97.72	-48.86	121.72

Table 4.21: Goodness of fit statistics for the ACL models with and without Proportional Odds respectively

Equation 4.5 and 4.6 show the multivariate ACL models without proportional odds assumptions for surgical stage 1 versus surgical stage 0 and surgical stage 2 versus surgical stage 1. Based on the goodness of fit statistics, these 2 ACL models without proportional odds were considered best in the classification of the surgical stage outcomes.

$$\begin{aligned} log[P(SS=0/SS=1)][= 2.499(0.622) - 0.0065(1.032)clinicalstage+ \\ 0.358(1.623)Vaginalinvolvement - 1.509(1.354) \\ Parametrialinvolvement - 1.261(0.657) \\ Symptom: Discharge - 1.261(0.657)Symptom: Pain] \end{aligned} \tag{4.5}$$

$$\begin{aligned} & log[P(SS=1/SS=2][= 0.994(1.044) - 2.394(1.258)clinicalstage \\ & - 1.420(1.552)Vaginalinvolvement - 1.403(1.444)Parametrialinvolvement \\ & + 0.327(1.027)Symptom: Discharge + 1.364(1.063)Symptom: Pain] \end{aligned} \tag{4.6}$$

The 3 ACL models with and without proportional odds were compared to determine the model best fit for the cervical cancer data. The multivariate ACL model with proportional odds had a misclassification rate of 32.00% and 37.32% whereas the multivariate ACL model without proportional odds had a misclassification rate of 29.33% and 37.03% when the train and validation datasets were utilized respectively. Clearly, there was an increase in misclassification by 5.32% and 7.70% respectively.

Confusion matrices 4.22 and 4.23 derived for the train and validation set from the multivariate ACL model without proportional odds gave an accuracy of 70.67%[95%CI : 59.02% − 80.62%] and 62.97%[95%CI : 62.01% − 63.92%] respectively. With reference to table 4.24, the sensitivity for surgical stage 0, 1 and 2 was 0.727, 0.500 and 0.571 for the train set and 0.667, 0.176 and 0.142 for the validation set. The specificity for surgical stage 0, 1 and 2 was $0.778, 0.808$ and 0.912 for the train set and $0.330, 0.797$ and for 0.870 for the validation set.

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	48	0	2
Surgical stage one	13	1	1
Surgical stage two	5	1	4

Table 4.22: Confusion matrix of the multivariate ACL model without proportional odds based on the train set.

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	6184	172	320
Surgical stage one	1884	45	90
Surgical stage two	1198	39	68

Table 4.23: Confusion matrix for the multivariate ACL model without proportional odds based on the validation set

	Sensitivity	Specificity	Prevalance
Train set			
Surgical stage zero	0.7273	0.7778	0.8800
Surgical stage one	0.5000	0.8082	0.0267
Surgical stage two	0.5714	0.9118	0.0933
Validation set			
Surgical stage zero	0.6674	0.3297	0.9266
Surgical stage one	0.1758	0.7974	0.0256
Surgical stage two	0.1423	0.8701	0.0478

Table 4.24: Sensitivity, Specificity and Prevalance for the multivariate ACL model without proportional odds based on the train and validation set respectively.

Table 4.25 shows the summary of the odds ratios for the ACL model without proportional odds. For the model 4.3 $log[P(SS = 0/SS = 1)]$ and model 4.4 $log[P(SS = 1/SS = 2)]$ respectively, for the patients diagnosed with FIGO clinical stage 2, the odds of being classified into surgical stage 1 versus surgical stage 0 was $0.99[0.13 - 7.51]$ times lower and the odds of being classified into surgical stage 2 versus surgical stage 1 was $0.09[0.01 - 1.07]$ times lower than for the patients with FIGO clinical stage 1 , holding all other predictors constant. Additionally, for the patients whose vaginal region was affected by the cervical cancer, the odds of being classified into surgical stage 1 versus surgical stage 0 was $1.43[0.06 - 34.47]$ times higher and the odds of being classified into surgical stage 2 versus surgical stage 1 was $0.24[0.01 - 5.06]$ times lower than for the patients without vaginal involvement, holding all other predictors constant. For the patients who had the parametrium affected by the cervical cancer, the odds of being classified under surgical stage 1 versus surgical stage 0 was $0.22[0.02 - 3.14]$ times lower and the odds of being classified under surgical stage 2 versus surgical stage 1 was $0.25[0.01 - 4.17]$ times lower than for patients without parametrial involvement, holding other predictors constant. For the patients with symptomatic vaginal discharge, the odds of being classified under surgical stage 1 versus surgical stage 0 was $0.28[0.08 - 1.03]$ times lower and the odds of being classified under surgical stage 2 versus surgical stage 1 was $1.39[0.19 - 10.39]$ times higher than for the patients without vaginal discharge, whilst holding other predictors constant. For the patients with symptomatic abdominal pain, the odds of being classified into surgical stage 1 versus surgical stage 0 was 0.30[0.08 − 1.09] times lower and the odds of being classified into surgical stage 1 versus surgical stage 0 was 3.91[0.49 − 31.42] times higher than for the patients without abdominal pain, whilst holding all other predictors constant.

Predictor variables	Odds ratio	CI : 2.5 %	CI: 97.5 %
(Intercept):1	12.17	3.60	41.20
(Intercept):2	2.70	0.35	20.91
Clinical.Stage 2:1	0.99	0.13	7.51
Clinical.Stage 2:2	0.09	0.01	1.07
Vaginal.Involvement Yes:1	1.43	0.06	34.47
Vaginal.Involvement Yes:2	0.24	0.01	5.06
Parametrial.Involvement Yes:1	0.22	0.02	3.14
Parametrial.Involvement Yes:2	0.25	0.01	4.17
Symptoms...Discharge Yes:1	0.28	0.08	1.03
Symptoms...Discharge Yes:2	1.39	0.19	10.39
Symptoms...Pain Yes:1	0.30	0.08	1.09
Symptoms...Pain Yes:2	3.91	0.49	31.42

Table 4.25: The Odds ratios for the ACL without proportional odds model

4.3 Regression Analysis based on the Predictive approach

The information from these patients was utilized to simulate a dataframe of 10000 patients with similar characteristics to the 75 patients. We were able to generate a random sample with a weighted sampling scheme by utilizing the sample() function in R studio.
Table 4.26 lists the counts for the train and validation (test) data sets respectively solely based on the simulated data.

Category	Train dataset	Validation dataset
Surgical stage 0	5341	1335
Surgical stage 1	1616	403
Surgical stage 2	1044	261

Table 4.26: Counts for patients in each surgical stage based on simulated data only

4.3.1 Polytomous (Multinomial) Logistic Regression Model

The reference category for the multinomial (polytomous) model was surgical stage 0. The 3 models fitted were the null model, univariate model and the multivariate model. We determined the presence of a relationship between the response variable and a combination of predictor variables by performing a likelihood ratio test. The results showed that the statistical significance of the probability of the model likelihood ratio (0.575) was 0.7502 indicating that there was no significant difference between the null (no predictors) and the univariate MLR model (FIGO clinical stage only). Moreover, the probability of the model likelihood ratio (7.834) was 0.4499 showing there was no significant difference between the univariate and the multivariate MLR models.

Table 4.27 shows the fitted univariate MLR model with the international FIGO clinical stage as the sole predictor variable. The output shows that the log odds of being in surgical stage 1 compared to surgical stage 0 (reference category) decreased by 0.031 if moving from clinical stage 1 to clinical stage 2 and the log odds of being in surgical stage 2 compared to surgical stage 0(reference category) increased by 0.043 if moving from clinical stage 1 to clinical stage 2. The intercepts show the log odds for surgical stage 0 (reference category).

Table 4.28 shows the summary for a multivariate model built with the inclusion of the 5 statistically significant predictors. The first model compares surgical stage 1 to surgical stage 0 (reference category)and the second model compares surgical stage 2 to surgical stage 0. We found that all the predictor variables were statistically insignificant and not associated to the surgical stage outcomes.

	Coefficient	Standard Errors	z-statistic	p-value
Surgical Stage 1				
Intercept	-1.189	0.032	-37.377	0.00000
Clinical Stage 2	-0.031	0.071	-0.444	0.65707
Surgical Stage 2				
Intercept	-1.641	0.038	-43.019	0.00000
Clinical Stage 2	0.043	0.083	0.517	0.60522

Table 4.27: Summary for a univariate MLR model built with the inclusion of the FIGO clinical stage only.

	Coefficient	Standard Errors	z-statistic	p-value
Surgical Stage 1				
Intercept	-1.202	0.049	-24.747	0.00000
Clinical Stage 2	-0.033	0.071	-0.462	0.64397
Vaginal Involvement: Yes	0.072	0.091	0.797	0.42568
Parametrial Involvement: Yes	-0.120	0.107	-1.125	0.26071
Symptoms Discharge: Yes	0.011	0.057	0.197	0.84363
Symptoms Pain: Yes	0.024	0.058	0.416	0.67733
Surgical Stage 2				
Intercept	-1.604	0.057	-28.025	0.00000
Clinical Stage 2	0.043	0.083	0.519	0.60391
Vaginal Involvement: Yes	0.155	0.105	1.477	0.13956
Parametrial Involvement: Yes	0.082	0.120	0.685	0.49357
Symptoms Discharge: Yes	-0.105	0.069	-1.519	0.12886
Symptoms Pain: Yes	-0.045	0.069	-0.653	0.51358

Table 4.28: Summary for a multivariate MLR model built with the inclusion of the 5 statistically significant predictors.

Table 4.7 displays the goodness of fit statistics for the 3 models. The full MLR model with the 5 statistically significant predictor variables had the highest AIC (with the lowest AIC being the null MLR model of 13743.34. The full MLR model gave the lowest residual deviance of 13754.93 with the highest log likelihood ratio of -6865.47. Thus, the multivariate multinomial logistic model was considered a better predictive fit for the simulated data compared to the univariate and null models.

	Deviance	Log Likelihood	AIC
Null MLR model	13739.34	-6869.67	13743.34
Univariate MLR model	13738.77	-6869.38	13746.77
Multivariate MLR model	13730.93	-6865.47	13754.93

Table 4.29: Goodness of fit statistics for the MLR models

Equation 4.7 and 4.8 show the multivariate MLR models for surgical stage 1 and surgical stage 2 both compared to the surgical stage 0 (the reference category). Based on the goodness-of-fit statistics, these 2 multivariate MLR models were considered best in the classification of the surgical stage outcomes.

$$\begin{aligned} log[P(SS=1|SS=0)][= -1.202(0.049) - 0.033(0.071)Clinicalstage+ \\ 0.072(0.091)Vaginalinvolvement - 0.120(0.107) \\ Parametrialinvolvement + 0.011(0.057) \\ Symptom: Discharge + 0.024(0.058)Symptom: Pain] \end{aligned} \tag{4.7}$$

$$
\begin{aligned}
&log[P(SS=2|SS=0][=-1.604(0.057)+0.043(0.083)clinicalstage\\
&+0.155(0.105)Vaginalinvolvement+0.082(0.120)Parametrialinvolvement \qquad (4.8)\\
&-0.105(0.069)Symptom:Discharge-0.045(0.069)Symptom:Pain]
\end{aligned}
$$

Confusion matrix 4.30 derived for the validation set gave an accuracy of 66.78%[95%CI : 64.67% − 68.85%] . With reference to table 4.31, the sensitivity for surgical stage 0, 1 and 2 are listed for the validation set.

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	1335	403	261
Surgical stage one	0	0	0
Surgical stage two	0	0	0

Table 4.30: Confusion matrix of the multivariate MLR model based on the validation set.

	Sensitivity	Specificity	Prevalance
Validation set			
Surgical stage zero	1.0000	0.0000	0.6678
Surgical stage one	0.0000	1.0000	0.2016
Surgical stage two	0.0000	1.0000	0.1306

Table 4.31: Sensitivity, specificity and prevalance for the multivariate MLR model based on the validation set.

Table 4.32 shows the odds ratios extracted from the MLR multivariate model.

	Odds ratio	2.5 %	97.5 %
(Intercept):1	0.30	0.27	0.33
(Intercept):2	0.20	0.18	0.22
Clinical.Stage 2:1	0.97	0.84	1.11
Clinical.Stage 2:2	1.04	0.89	1.23
Vaginal.Involvement Yes:1	1.07	0.90	1.28
Vaginal.Involvement Yes:2	1.17	0.95	1.43
Parametrial.Involvement Yes:1	0.89	0.72	1.09
Parametrial.Involvement Yes:2	1.09	0.86	1.37
Symptoms...Discharge Yes:1	1.01	0.90	1.13
Symptoms...Discharge Yes:2	0.90	0.79	1.03
Symptoms...Pain Yes:1	1.02	0.92	1.15
Symptoms...Pain Yes:2	0.96	0.84	1.09

Table 4.32: The odds ratios for the MLR multivariate model were extracted.

With the best fit model being the multivariate MLR model, the odds ratios were extracted from the models. The odds ratios 4.32 extracted from the multivariate multinomial logistic model which displayed the best fit model for the simulated data. The odds of being classified into surgical stage 1 over surgical stage 0 was 0.97 [CI: $0.84 - 1.11$] lower for patients diagnosed with FIGO clinical stage 2 versus those diagnosed with FIGO clinical stage 1 while holding all other predictors constant. Moreover, the odds of being classified into surgical stage 2 over surgical stage 0 was 1.04 CI:[$0.89 - 1.23$] times higher for patients diagnosed with FIGO clinical stage 2 versus those diagnosed with FIGO clinical stage 1 while holding all other predictors constant. The odds of being classified into surgical stage 1 over surgical stage 0 was 1.07 CI:[$0.90 - 1.28$] times higher and the odds of being in surgical stage 2 over surgical stage 0 was1.17 CI: [$0.95 - 1.43$] times higher for patients with the vaginal region observed to be affected by the cancer during diagnosis versus those without any vaginal involvement while holding all other predictors constant. The odds of being classified into surgical stage 1 over surgical stage 0 was 0.89 CI: [$0.72 - 1.09$] times lower and the odds of being classified into surgical stage 2 over surgical stage 0 was 1.09 CI: [$0.86 - 1.37$] times higher for patients with the parametrial region affected by the cancer versus those without any parametrial involvement while holding other predictors constant. The odds of being classified into surgical stage 1 over surgical stage 0 was 1.01 CI: [$0.90 - 1.13$] times higher and the odds of being classified into surgical stage 2 over surgical stage 0 was 0.90 CI: [$0.79 - 1.03$] times lower for the patients with symptomatic vaginal discharge during diagnosis versus those without the symptomatic vaginal discharge, holding all other predictors constant. The odds of being classified into surgical stage 1 over surgical stage 0 was 1.02 CI:[$0.92 - 1.15$] times higher and in contrast, the odds of being into surgical stage 2 over surgical stage 0 was 0.96 CI:[$0.84 - 1.09$] times lower for the patients with symptomatic lower abdominal pain versus those without any pain, holding all other predictors constant.

4.3.2 Continuation Ratio Model

Continuation Ratio Model with and without Proportional Odds

The tables below show the output for the CR null, univariate and multivariate models with and without proportional odds respectively. The univariate CR model with and without proportional odds output considered the FIGO clinical stage as the only predictor. Clearly, the CR univariate model with proportional odds showed that the FIGO clinical stage had no significant effect on the surgical stage response with a p-value of 0.9212 when the regression coefficient was common. The estimated log regression coefficient for the FIGO clinical stage , $\beta = 0.004(0.038), z = 0.099$ and $p = 0.9212$ showed that the FIGO clinical stage upon diagnosis had a positive effect on the surgical stage responses. The CR

model without proportional odds gave separate effects. All the 5 predictors were found not to be statistically significant. For having a surgical stage outcome greater than surgical stage 1 relative to being in surgical stage 1, the estimated log regression coefficient for the FIGO clinical stage was $\beta = 0.027(0.049)$, z-value= 0.555 and a p-value of 0.579 indicating that the FIGO clinical stage had no significant effect on the surgical stage 1 responses. In addition, having a surgical stage outcome greater than surgical stage 2 relative to being in surgical stage 2, the estimated logit regression coefficient for FIGO clinical stage was $\beta = -0.030(0.058)$, z-value=-0.517 and a p-value of 0.605 indicating that the FIGO clinical stage had no significant effect on surgical stage 2 responses.
The multivariate CR model with and without proportional odds output with the 5 predictors that were found to be statistically significant were included.For the CR multivariate model with proportional odds, the FIGO clinical stage estimated log regression coefficient, $\beta = 0.004(0.038)$, z-value=0.116 had a positive effect on the surgical stage responses. In contrast, a p-value of 0.9079 showed that the FIGO clinical stage is not a statistically significant predictor for having a surgical stage outcome greater than surgical stage 1 relative to being in surgical stage 1 as well as having a surgical stage outcome greater than surgical stage 2 relative to being in surgical stage 2. The remaining 4 predictors that were not statistically significant to the surgical stage responses were the vaginal involvement, the parametrial involvement, symptomatic vaginal discharge and lower abdominal pain. For the CR multivariate model without proportional odds, we got separate effects for the surgical stage responses.For having a surgical stage outcome greater than surgical stage 1 relative to being in surgical stage 1, the FIGO clinical stage and the parametrial involvement had no effect on the surgical stage outcome with an estimated logit regression coefficient of $\beta = 0.028(0.049)$, z-value=0.574 and p-value of 0.5660 and an estimated logit regression coefficient of $\beta = 0.134(0.105)$, z-value=1.277 and p-value of 0.2017 respectively. The estimated logit coefficients for the vaginal involvement,symptomatic discharge and symptomatic lower abdominal pain were not statistically significant with estimated logit regression coefficients of $\beta = -0.046(0.089)$, z-value=-0.514 and p-value of 0.6073, $\beta = -0.028(0.056)$, z-value=-0.503 and p-value of 0.6150 and $\beta = -0.031(0.056)$, z-value=-0.554 and p-value of 0.5799 . Also, it was clear that for having a surgical stage outcome greater than surgical stage 2 relative to being in surgical stage 2, the FIGO clinical stage had a no effect and was not statistically significant with an estimated logit regression coefficient of $\beta = -0.030(0.058)$, z-value=-0.513 and p-value of 0.0608. The estimated logit regression coefficients for vaginal involvement,parametrial involvement, symptomatic vaginal discharge and symptomatic lower abdominal pain were not statistically significant with no effect on the surgical stage responses and were found to be $\beta = -0.156(0.105)$,

z-value=-1.490 and p-value of 0.1362 ,$\beta = -0.084(0.119)$, z-value=-0.701 and p-value of 0.4835,$\beta = 0.106(0.069)$, z-value=1.530 and p-value of 0.1261 and $\beta = 0.046(0.069)$, z-value=0.666 and p-value of 0.5055.

	Coefficient	Standard Errors	z-statistic	p-value
CR Model with Proportional Odds				
Intercept1	1.376	0.032	43.120	0.00000
Intercept2	1.634	0.037	43.870	0.0000
Clinical Stage 2	0.004	0.038	0.099	0.92117
CR Model without Proportional Odds				
Intercept1	1.385	0.035	40.063	0.00000
Intercept2	1.620	0.041	39.226	0.00000
Clinical Stage 1	0.027	0.049	0.555	0.579
Clinical Stage 2	-0.030	0.058	-0.517	0.605

Table 4.33: The Summary for a CR univariate model with the inclusion of the FIGO clinical stage predictor.

	Coefficient	Standard Errors	z-statistic	p-value
CR Model with Proportional Odds				
Intercept1	1.372	0.042	32.565	0.00000
Intercept2	1.630	0.046	35.200	0.00000
Clinical Stage 2	0.004	0.038	0.116	0.90789
Vaginal Involvement: Yes	-0.090	0.068	-1.333	0.18248
Parametrial Involvement: Yes	0.043	0.079	0.541	0.58828
Symptoms Discharge: Yes	0.025	0.044	0.580	0.56219
Symptoms Pain: Yes	0.000	0.044	-0.011	0.99160
CR Model without Proportional Odds				
Intercept1	1.405	0.050	28.170	0.00000
Intercept2	1.583	0.059	26.671	0.00000
Clinical Stage 1	0.028	0.049	0.574	0.56600
Clinical Stage 2	-0.030	0.058	-0.513	0.060794
Vaginal Involvement: Yes1	-0.046	0.089	-0.514	0.60731
Vaginal Involvement: Yes2	-0.156	0.105	-1.490	0.13616
Parametrial Involvement: Yes1	0.134	0.105	1.277	0.20165
Parametrial Involvement: Yes2	-0.084	0.119	-0.701	0.48352
Symptoms Discharge: Yes1	-0.028	0.056	-0.503	0.61504
Symptoms Discharge: Yes2	0.106	0.069	1.530	0.12605
Symptoms Pain: Yes1	-0.031	0.056	-0.554	0.57992
Symptoms Pain: Yes2	0.046	0.069	0.666	0.50551

Table 4.34: The Summary for a CR multivariate model with the inclusion of the 5 predictors for Surgical Stage 1 and Stage 2.

Table 4.35 and 4.36 show the goodness of fit statistics for the 2 CR multivariate models (with and without proportional odds). An AIC of 13743.34 shows that the null CR models would be a good fit. However, a low residual deviance and likelihood ratio of 13730.89 and −6865.44 respectively show that the multivariate CR model without proportional odds is a better fit. Moreover, the VGAM likelihood ratio test was carried out for the 2 CR multivariate models and a chi-square p-value of 0.303 showed that the fits were not significantly different and thus, the multivariate CR model without proportional odds was found to be adequate.

	Deviance	Log Likelihood	AIC
Null CR model	13739.34	-6869.67	13743.34
Univarite CR model	13739.33	-6869.67	13745.33
Multivariate CR model	13736.92	-6868.46	13750.92

Table 4.35: Goodness of fit statistics for the CR models with Proportional Odds

	Deviance	Log Likelihood	AIC
Null CR model	13739.34	-6869.67	13743.34
Univariate CR model	13738.77	-6869.38	13746.77
Multivariate CR model	13730.89	-6865.44	13754.89

Table 4.36: Goodness of fit statistics for the CR models without Proportional Odds

The multivariate CR model with proportional odds did give better goodness of fit statistics. However, all the 5 predictors were found not to be statistically significant to the prediction of the surgical stage.A log likelihood ratio test for the null and univariate CR models with PPOs versus the full CR model with PPOs each gave a p-value of 0.788 and 0.660 respectively showing that their was no significant differences among the 3 models.
Moreover, the log likelihood ratio test for the null and univariate CR models without PPOs versus the full CR model without PPOs each gave a p-value of 0.584 and 0.445 respectively showing that their was no significant differences among the 3 models.

Below is the confusion matrix 4.3.2 for the multivariate CR model without PPOs. The model gave an accuracy of 66.78% [CI: 64.67% - 68.85%].

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	1335	0	0
Surgical stage one	403	0	0
Surgical stage two	261	0	0

Table 4.37: Confusion matrix of the multivariate CR model without proportional odds based on the validation set.

Equation 4.9 and 4.10 show the multivariate CR model without proportional odds assumptions for surgical stage 1 and surgical stage 2. Based on the goodness-of-fit statistics, these 2 CR models without proportional odds were considered best in the classification of the surgical stage outcomes.

$$\begin{aligned} log[P(SS=1|SS\geq 1)][= &1.405(0.050)+0.028(0.049)clinicalstage-\\ &0.046(0.089)Vaginalinvolvement+0.134(0.105)\\ &Parametrialinvolvement-0.028(0.056)\\ &Symptom:Discharge-0.031(0.056)Symptom:Pain] \end{aligned} \tag{4.9}$$

$$\begin{aligned} log[P(SS=2|SS\geq 2)][= &1.583(0.059)-0.030(0.058)clinicalstage-\\ &0.156(0.105)Vaginalinvolvement-0.084(0.119)\\ &Parametrialinvolvement+0.106(0.069)\\ &Symptom:Discharge+0.046(0.069)Symptom:Pain] \end{aligned} \tag{4.10}$$

	Odds ratio	2.5 %	97.5 %
(Intercept):1	4.076	3.697	4.495
(Intercept):2	4.867	4.333	5.467
Clinical.Stage 2:1	1.028	0.934	1.132
Clinical.Stage 2:2	0.970	0.865	1.088
Vaginal.Involvement Yes:1	0.955	0.803	1.137
Vaginal.Involvement Yes:2	0.855	0.696	1.051
Parametrial.Involvement Yes:1	1.144	0.931	1.405
Parametrial.Involvement Yes:2	0.920	0.728	1.162
Symptoms...Discharge Yes:1	0.972	0.870	1.086
Symptoms...Discharge Yes:2	1.111	0.971	1.272
Symptoms...Pain Yes:1	0.969	0.868	1.083
Symptoms...Pain Yes:2	1.047	0.915	1.198

Table 4.38: The odds ratios for the CR multivariate model without proportional odds were extracted.

By utilizing the multiclass.roc() function, we found that the multi-class area under the curve for the univariate and multivariate CR models without PPOs were 0.4996 and 0.5071 respectively

.For the model 4.9 $logit[P(SS = 1|SS \geq 1]$ and model 4.10 $logit[P(SS = 2|SS \geq 2]$ respectively, the odds of having an outcome greater than surgical stage 1 relative to being in surgical stage 1 was $1.028[0.93 - 1.13]$ times higher and the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $0.97[0.87 - 1.09]$ times lower among the patients diagnosed with FIGO clinical stage 2 compared to the patients diagnosed with FIGO clinical stage 1, after controlling for the effects of other predictors in the model. In addition, the odds of having an outcome greater than surgical stage 1 relative to being in surgical stage 1 was $0.96[0.80 - 1.14]$ times lower and the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $0.86[0.70 - 1.05]$ times lower among the patients considered to have the vaginal region affected by the cancer (vaginal involvement) compared to the patients without any vaginal involvement after controlling for the effects of other predictors in the model. The odds of having an outcome greater than surgical stage 1 relative to being in surgical stage 1 was $1.14[0.93 - 1.41]$ times higher and the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $0.92[0.73 - 1.16]$ times lower among the patients considered to have the parametrium region affected by the cervical cancer (parametrial involvement) compared to the patients without any parametrial involvement after controlling for the effects of other predictors in the model. The odds of having an outcome greater that surgical stage 1 relative to being in surgical stage 1 was $0.97[0.87 - 1.09]$ times lower and in contrast, the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $1.11[0.97 - 1.27]$ times higher among the patients with symptomatic vaginal discharge (Symptoms: Discharge) compared to the patients who did not have symptomatic vaginal discharge after controlling for the effects of other predictors in the model. The odds of having an outcome greater than surgical stage 1 relative to being in surgical stage 1 was $0.97[0.87 - 1.08]$ times lower and in contrast, the odds of having an outcome greater than surgical stage 2 relative to being in surgical stage 2 was $1.05[0.92 - 1.20]$ times higher among the patients displaying symptomatic lower abdominal pain (Symptoms: Pain) compared to the patients without symptomatic lower abdominal pain after controlling for the effects of other predictors in the model.

4.3.3 Adjacent Category Logistic Regression Model

Adjacent Category Logistic Regression Model with and without Proportional Odds

The tables below shows the summary of the ACL null, univariate and multivariate models with and without proportional odds. For the ACL univariate model with proportional odds, we found that all the 5 predictors were not statistically significant. The FIGO clinical stage had no statistically significant effect on the surgical stage response with a p-value of 0.69. The estimated logit regression coefficient for the FIGO clinical stage, $\beta = 0.015(0.039)$, z-value=0.387 and a p-value 0.69 which showed that the FIGO clinical stage upon diagnosis had no effect on each adjacent surgical stage response category (common regression coefficients). The ACL univariate model without proportional odds gave separate effects. We found that the estimated logit regression coefficient $\beta = 0.030(0.070)$, z-value=-0.433 and the p-value=0.665 indicated that the log odds of being in surgical stage 1 versus surgical stage 0 was 0.030 when the FIGO clinical stage increased by 1 unit, holding all other predictors constant. In addition, the estimated logit regression coefficient $\beta = -0.009(0.098)$, z-value=-0.090 and the p-value=0.928 indicated that the log odds of being in surgical stage 2 versus surgical stage 1 was -0.009 when the FIGO clinical stage increased by 1 unit, holding all other predictors constant. However, the FIGO clinical stage had no significant effect on the probability of being in surgical stage 1 versus surgical stage 0 and on the probability of being in surgical stage 2 versus surgical stage 1 . For the ACL multivariate model with proportional odds , the 5 predictors were not statistically significant to the surgical stage responses. For the ACL multivariate model without proportional odds, we get separate effects for the surgical stage responses. The FIGO clinical stage had no effect on the probability of being classified under surgical stage 1 versus surgical stage 0 and on the probability of being classified under surgical stage 2 versus surgical stage 1 with the estimated logit regression coefficient $\beta = 0.031(0.070)$, z-value = 0.435 and a p-value of 0.664 indicating that it is not a significant predictor and the log-odds of being classified under surgical stage 2 versus surgical stage 1 being $-0.011(0.098)$,z-value = -0.107 and a p-value of 0.915 when the FIGO clinical stage increased by 1 unit, holding all other predictors constant. Moreover, the remaining 4 predictor variables had no statistical significance to the probability of being classified either under surgical stage 1 versus surgical stage 0 or surgical stage 2 versus surgical stage 1.

	Coefficient	Standard Errors	z-statistic	p-value
ACL Model with Proportional Odds				
Intercept1	1.196	0.029	42.11	$< 2e$ - 16
Intercept2	0.437	0.040	11.00	$< 2e$ - 16
ACL Model without Proportional Odds				
Intercept1	1.196	0.029	42.11	$< 2e$ - 16
Intercept2	0.437	0.040	11.00	$< 2e$ - 16

Table 4.39: The Summary for a ACL null model.

	Coefficient	Standard Errors	z-statistic	p-value
ACL Model with Proportional Odds				
Intercept1	1.192	0.030	40.396	$< 2e$ - 16
Intercept2	0.434	0.040	10.716	$< 2e$ - 16
Clinical Stage 2	0.015	0.039	0.387	0.69
ACL Model without Proportional Odds				
Intercept1	1.189	0.032	37.308	$< 2e$ - 16
Intercept2	0.439	0.045	9.843	$< 2e$ - 16
Clinical Stage 1	0.030	0.070	0.433	0.665
Clinical Stage 2	-0.009	0.098	-0.090	0.928

Table 4.40: The Summary for a ACL univariate model with the inclusion of the FIGO clinical stage predictor.

	Coefficient	Standard Errors	z-statistic	p-value
ACL Model with Proportional Odds				
Intercept1	1.181	0.036	33.196	$< 2e$ - 16
Intercept2	0.424	0.045	9.425	$< 2e$ - 16
Clinical Stage 2	0.014	0.039	0.372	0.710
Vaginal Involvement: Yes	-0.059	0.050	-1.181	0.238
Parametrial Involvement: Yes	-0.075	0.056	-1.341	0.180
Symptoms Discharge: Yes	0.042	0.032	1.330	0.183
Symptoms Pain: Yes	0.014	0.032	0.430	0.667
ACL Model without Proportional Odds				
Intercept1	1.181	0.048	24.539	$< 2e$ - 16
Intercept2	0.425	0.067	6.337	$2.35e$ - 10
Clinical Stage 1	0.031	0.070	0.435	0.6637
Clinical Stage 2	-0.011	0.098	-0.107	0.9149
Vaginal Involvement: Yes1	-0.053	0.092	-0.577	0.5642
Vaginal Involvement: Yes2	-0.068	0.126	-0.540	0.5893
Parametrial Involvement: Yes1	0.037	0.106	0.353	0.7241
Parametrial Involvement: Yes2	-0.239	0.141	-1.698	0.0895
Symptoms Discharge: Yes1	0.020	0.058	0.353	0.7240
Symptoms Discharge: Yes2	0.077	0.081	0.946	0.3440
Symptoms Pain: Yes1	0.005	0.058	0.081	0.9351
Symptoms Pain: Yes2	0.028	0.081	0.342	0.7325

Table 4.41: The Summary for a ACL multivariate model with the inclusion of the 5 predictors for Surgical Stage 1 and Stage 2.

Tables 4.42 and 4.43 display the goodness of fit statistics for the 3 ACL multivariate models with and without proportional odds. An AIC of 13743.34 indicates that the null ACL models would be a good fit. However, the multivariate ACL without proportional odds had the lowest residual deviance and the highest log likelihood ratio of 13732.25 and -6866.13 respectively. We carried out a VGAM likelihood ratio test for the 2 multivariate ACL models and a chi-square p-value of 0.865 indicating that both fits were not significantly different from each other.

	Deviance	Log Likelihood	AIC
Null ACR model	13739.34	-6869.67	13743.34
Univariate ACL model	13739.19	-6869.60	13745.19
Multivariate ACL model	13734.14	-6867.07	13748.14

Table 4.42: Goodness of fit statistics for the ACL models with Proportional Odds

	Deviance	Log Likelihood	AIC
Null ACL model	13739.34	-6869.67	13743.34
Univariate ACL model	13739.12	-6869.56	13747.12
Multivariate ACL model	13732.25	-6866.13	13756.25

Table 4.43: Goodness of fit statistics for the ACL models without Proportional Odds

The 3 ACL models with and without proportional odds were compared to determine the model best fit for the simulated data. The multivariate ACL model with proportional odds had a misclassification rate of 33.22% and the multivariate ACL model without proportional odds had a misclassification rate of 33.22% when the validation datasets were utilized.
Confusion matrices 4.3.3 derived for the validation set from the multivariate ACL model without proportional odds gave an accuracy of $66.78\%[95\%CI : 64.67\% - 68.85\%]$.

	Actual		
	Surgical stage zero	Surgical stage one	Surgical stage two
Prediction			
Surgical stage zero	1335	0	0
Surgical stage one	403	0	0
Surgical stage two	261	0	0

Table 4.44: Confusion matrix of the multivariate ACL model without proportional odds based on the validation set.

Equation 4.11 and 4.12 show the multivariate ACL model without proportional odds for surgical stage 1 versus surgical stage 0 and surgical stage 2 versus surgical stage 1.

$$\begin{aligned} log[P(SS = 0|SS = 1)][= 1.181(0.048) + 0.031(0.070)clinicalstage- \\ 0.053(0.092)Vaginalinvolvement + 0.037(0.106) \\ Parametrialinvolvement + 0.020(0.058) \\ Symptom : Discharge + 0.005(0.058)Symptom : Pain] \end{aligned} \tag{4.11}$$

$$\begin{aligned} &log[P(SS = 1|SS = 2][= 0.425(0.067) - 0.011(0.098)clinicalstage \\ &- 0.068(0.126)Vaginalinvolvement - 0.239(0.141)Parametrialinvolvement \\ &+ 0.077(0.081)Symptom : Discharge + 0.028(0.081)Symptom : Pain] \end{aligned} \tag{4.12}$$

Table 4.45 shows the summary of the odds ratios for the ACL model without proportional odds based on the simulated data. For the model 4.11 $log[P(SS = 0/SS = 1)]$ and model 4.12 $log[P(SS = 1/SS = 2)]$ respectively, for the patients diagnosed with FIGO clinical stage 2, the odds of being classified into surgical stage 1 versus surgical stage 0 was $1.03[0.90 - 1.18]$ times higher and the odds of being classified into surgical stage 2 versus surgical stage 1 was $0.99[0.82 - 1.20]$ times lower than for the patients with FIGO clinical stage 1 , holding all other predictors constant. Additionally, for the patients whose vaginal region was affected by the cervical cancer, the odds of being classified into surgical stage 1 versus surgical stage

0 was 0.95[0.79−1.14] times lower and the odds of being classified into surgical stage 2 versus surgical stage 1 was 0.93[0.73−1.20] times lower than for the patients without vaginal involvement, holding all other predictors constant. For the patients who had the parametrium affected by the cervical cancer, the odds of being classified under surgical stage 1 versus surgical stage 0 was 1.04[0.84 − 1.28] times higher and the odds of being classified under surgical stage 2 versus surgical stage 1 was 0.79[0.60−1.04] times lower than for patients without parametrial involvement, holding other predictors constant. For the patients with symptomatic vaginal discharge, the odds of being classified under surgical stage 1 versus surgical stage 0 was 1.02[0.91 − 1.14] times higher and the odds of being classified under surgical stage 2 versus surgical stage 1 was 1.08[0.92 − 1.27] times higher than for the patients without vaginal discharge, whilst holding other predictors constant. For the patients with symptomatic abdominal pain, the odds of being classified into surgical stage 1 versus surgical stage 0 was 1.00[0.90 − 1.12] times higher and the odds of being classified into surgical stage 1 versus surgical stage 0 was 1.03[0.88−1.20] times higher than for the patients without abdominal pain, whilst holding all other predictors constant.

Predictor variables	Odds ratio	2.5 %	97.5 %
(Intercept):1	3.26	2.97	3.58
(Intercept):2	1.53	1.34	1.74
Clinical.S2:1	1.03	0.90	1.18
Clinical.S2:2	0.99	0.82	1.20
Vaginal.InvolvementYes:1	0.95	0.79	1.14
Vaginal.InvolvementYes:2	0.93	0.73	1.20
Parametrial.InvolvementYes:1	1.04	0.84	1.28
Parametrial.InvolvementYes:2	0.79	0.60	1.04
Symptoms...DischargeYes:1	1.02	0.91	1.14
Symptoms...DischargeYes:2	1.08	0.92	1.27
Symptoms...PainYes:1	1.00	0.90	1.12
Symptoms...PainYes:2	1.03	0.88	1.20

Table 4.45: The Odds ratios for the ACL without proportional odds model

By utilizing the multiclass.roc() function, we found that the multi-class area under the curve for the multivariate ACL model without PPOs was 0.5214 .

Table 4.46 shows the values of the area under the curve (AUC) for the best fits of each model. In his article, Lans (2019) describes the AUC metric as the probability that an observation with a positive class will have a greater predicted probability than an observation in a negative class. The AUC is considered similar to the R-square value in logistic regression. The AUC results are within 0.5 indicating that the 3 models were unable to discriminate between the surgical stage categories.

Ordinal Multivariate Model	Multi-class ROC (area under the curve)
Multinomial Logistic Regression model	0.506
Continuation Ratio without Proportional Odds Model	0.507
Adjacent Category Logistic without proportional odds model	0.521

Table 4.46: The ROC (Area under the curve) for the 3 best fit models.

Chapter 5

Discussion

5.1 Introduction

Ordinal techniques utilize the ordering of the levels and do not assume equal spacing between the levels of the response variable as stated by Meisner et al. (2018). The World Health Organization stated that the availability of effective treatment amongst women $30 - 49$ years of age is not being captured. Still, the aim of cervical cancer screening is to detect the pre-cancerous changes on the cervix which may lead to cancer. The objective of this study was to evaluate the predictive performance of 3 regression models for ordinal responses on the surgical stage of women treated surgically for invasive cervical cancer based on an inferential and predictive approach. The inferential approach relied on utilizing the same data and in contrast, the predictive approach required the use of simulated data with similar characteristics to the original data.

The results may provide an understanding of the future possibilities of using predictive algorithms in the Kenyan oncology setting and furthermore, allow for informed based surgical treatment patient selection. In our study, a bivariate analysis was carried out to determine the factors associated with the surgical stage after surgical treatment of cervical cancer. Thereafter, the relationship between the surgical stage outcomes and 5 statistically significant predictor variables were investigated by developing, validating and comparing the odds ratios of 3 types of regression models for ordinal responses. The relationships between the surgical stage and 5 statistically significant variables were investigated by applying regression models and comparing the odds ratios.The main point of interest in this discussion was to assess the benefit and possibility of utilizing ordinal regression models.

5.2 Application to surgically-treated cervical cancer patients

Inferential Approach

It is important to understand that the main aim of surgical treatment for early stage cervical cancer is to extract all the cancer before the cancer spreads to the lymph nodes and tissues around the tumour. Successful surgical treatment outcomes result in cost effectiveness and proper utilization of resources by avoiding the need for alternative therapy. Thus, consider the question: "Among the 3 regression models for ordinal responses, which model best estimates and predicts the surgical stage prior to surgical treatment of cervical cancer?" The results indicated that the FIGO clinical stage, parametrial involvement, vaginal involvement, symptomatic vaginal discharge and lower abdominal pains were statistically significant and independently associated with the surgical stage and could be taken into account when determining whether surgery would be the most effective mode of treatment therapy to be assigned to a patient with confirmed cervical cancer. Such would considerably increase the efficiency and effectiveness of the surgical treatment of cervical cancer.In addition, the findings from this study suggest that among the 3 ordinal regression models, the continuation ratio model without proportional odds was found to best classify the surgical stages of the patients with a misclassification rate of 33.22% . Although the 3 models are similar in that they fit multiple simultaneous binary logits, there is some restructuring of categories. The CR model fits 2 logits on each consecutive step; in terms of dummy variables, with the increasing '0' category, the '1' category is considered the higher category. The MLR model compares each of the surgical stages 1 to 2 with surgical stage 0 (the reference category) in 2 simultaneous logit models and the ACL model fits logit models to 2 adjacent pairs of surgical stage categories.

The results showed that the inferential approach focused on statistical significance of the FIGO clinical stage on women classified under surgical stage 2. However, the small sample size of the train set could have brought about bias and uncertainty on the results.

The results further showed that for each model, the multivariate models took precedence which indicated that a combination of predictors could best determine the surgical stage outcome of a patient prior to surgery.

The multivariate CR model without proportional odds presented an AIC value of 118.89 indicating that it would be the best model to select for the cervical cancer data. The study demonstrated a similarity between the MLR and ACL model.

The multivariate MLR and ACL model without proportional odds had similar likelihood ratios of -48.86 whilst the CR model without proportional odds had a likelihood ratio of -47.44 showing that the later model was statistically different from the 2 models. In terms of model selection, the multivariate MLR and ACL model without proportional odds presented similar AIC values of 121.72. de Rooij (2019) stated that when β_j is utilized instead of a single β, then the ACL model becomes equivalent to the MLR model. Also, the similarity in AIC statistic values could be attributed to the small sample size of the cervical cancer data. This fact could explain the similarity in the goodness-of-fit statistics between the MLR and ACL models. In order to test the proportional odds assumption, the CR model and the ACL models were fitted with and without proportional odds. The accuracy of the CR model with and without proportional odds was 70.67% and 69.33% respectively. However, the goodness-of-fit statistics showed that the CR model without proportional odds gave the lowest deviance of 94.89 and a low AIC statistic of 118.72. On analyzing the results, the CR null models with and without proportional odds gave similar coefficients and negligible differences were observed. The univariate and multivariate CR models without proportional odds gave separate effects for each independent variable. Both univariate CR models supported that the FIGO clinical stage did have a significant positive influence on the surgical stage outcomes. Although the CR model without proportional odds gave the lowest deviance and a low AIC statistic, this particular model showed that information on the FIGO clinical stage had a higher influence on the patients with surgical stage 2 compared to those with FIGO clinical stage 1.

We compared the odds ratios of the 3 multivariate models. The odds ratio, according to Norton and Dowd (2018), is not an absolute number. In addition, Norton and Dowd (2018) asserted that odds ratios are simple to compute and can be applied to discrete and continuous explanatory variables. The odds ratios, as Szumilas (2010) explained in her article, compare the relative odds of the response (in our case, surgical stage), given exposure to explanatory variables of interest. She further expounds that the odds ratios can ascertain whether a particular exposure is a risk factor and compare the magnitude of various risk factors for the specific response. The 95% odds ratios confidence intervals estimate the precision of the odds ratios and are considered a substitute for the presence of statistical significance when the null value (OR=1) is not overlapped. Low levels of precision are indicated by large confidence intervals whereas high levels of precision are indicated by small confidence intervals.

The FIGO clinical stage had a higher influence on women whose odds of having a surgical stage greater than surgical stage 2 relative to being in surgical stage 2 (see

Table 4.17). Though the results gave large confidence intervals indicative of low precision, a statistically significant p-value (0.0349) and confidence intervals that did not span the null value (OR=1) confirmed the result. The OR for the other 4 predictors showed decreased odds of having a surgical stage greater than surgical stage 2 relative to being in surgical stage 2. Also, there was decreased odds of having a surgical stage greater than surgical stage 1 relative to being in surgical stage 1 with the confidence intervals for the 5 statistically significant predictors spanning the null value (OR=1). Clearly, there was no statistical significance with the regression coefficients having p-values at > 0.05. The likelihood chi-square ratio test showed that the CR model without proportional odds (chi-square p-value=0.0823) is adequate compared to the CR model with proportional odds.

The utilization of ordinal regression models enables significant independent variables to be identified with their effect on the classification of patients into surgical stages. The MLR model was utilized so as to compared the results when the ordinal scales were assumed nominal and the surgical stage outcomes analyzed as binary logistic regression. Based on the results from the univariate and multivariate MLR models, the findings suggest that the FIGO clinical stage alone could not adequately predict the surgical stage outcomes based on the p-values being > 0.05 (see Table 4.3 and 4.4. By utilizing the multinom function, Rodríguez (2019a) emphasized its simplicity and the fact that the function works well with vector outcomes and factor predictor variables. The multinom function chose the first response category as the reference category which was surgical stage 0 in our case. The MLR model was fitted in which the probability of each outcome was estimated in one maximum likelihood procedure. Our findings concerning the FIGO clinical stage in determining the surgical stage of a cervical cancer patient was not sufficient enough to determine the surgical treatment option. With reference to table 4.3, the univariate MLR model had a logistic coefficient of 2.66 showing that the FIGO clinical stage had more of a higher influence on determining the classification into surgical stage 2 with a less and negative influence noted on classification into surgical stage 1 which had a logistic coefficient of -0.06 both being compared to the baseline category (surgical stage 0). This is supported by a graphical effects display for the surgical stage which apparently showed that the patients classified under surgical stage 2 were more likely to have FIGO clinical stage 2 cervical cancer compared to FIGO clinical stage 1 and patients classified under surgical stages 0 and 1 were less likely to have FIGO clinical stage 2 cervical cancer compared to FIGO clinical stage 1. Our findings were complimented by Gold et al. (2008) who stated that clinical staging is only 60% accurate. However, the predictive approach demonstrated that the 5 predictors were not statistically significant in predicting the surgical stage

outcomes.
In a study aimed at ascertaining the presence, distribution and metastatic involvement of parametrial lymph nodes of patients undergoing radical hysterectomy, Girardi et al. (1989) found a linear association between both the FIGO clinical stage, tumor volume and the frequency of positive parametrial lymph nodes. Moreover, with positive lymph nodes, the recurrence rate increased and the survival rate dropped. The multivariate MLR model findings completely ruled out the FIGO clinical stage as having any influence on classifying a patient under surgical stage 1 (p-value=0.99441) versus surgical stage 0 and in contrast having positive and significant influence on classifying a patient under surgical stage 2 (p-value=0.02297) versus surgical stage 0. Although the presence of parametrial involvement was not statistically significant, there were higher odds of classifying a patient under surgical stage 2 versus surgical stage 0 when the FIGO clinical stage was found to be statistically significant to predicting the surgical stage outcome. Based on the inferential approach, FIGO clinical stage took precedence as the key independent predictor. Unlike Girardi et al. (1989), the data lacked sufficient information on tumor size and frequency of positive parametrial lymph nodes; rather, our study had a binary outcome (yes/no) for the cancer being involved with the parametrium.

The inferential approach conveyed that the ACL multivariate model without proportional odds had the highest accuracy (70.67%[59.02% − 80.62%]). Based on the inferential approach, both the univariate ACL models with and without proportional odds support that the FIGO clinical stage has a significant effect on the surgical stage outcomes. Based on the goodness-of-fit statistics, the multivariate ACL model without proportional odds had the lowest deviance and the highest log-likelihood ratio. It is intriguing to observe that the 5 predictors were statistically insignificant on the probability of being classified under surgical stage 1 versus surgical stage 0 and on the probability of being classified under surgical stage 2 versus surgical stage 1.The FIGO clinical stage showed greater influence on surgical stage outcomes in the ACL model with proportional odds. Based on the multivariate ACL model without proportional odds, there was decreased odds of being classified under surgical stage 1 versus surgical stage 0 among the women with FIGO clinical stage 2 compared to those with FIGO clinical stage 1 with the OR confidence intervals displaying statistical non-significance. In the case of being classified under surgical stage 1 versus surgical stage 0, we noted decreased odds for all the predictors with the exception of the presence of vaginal involvement (OR=1.43[0.06 − 34.47]) with low precision and statistical non-significance. Similarly, we make out decreased odds of being classified under surgical stage 2 versus surgical stage 1 among the women with

FIGO clinical stage 2 compared to those with FIGO clinical stage 1 with the OR confidence intervals displaying statistical non-significance. We noted that decreased odds of being classified under surgical stage 2 versus surgical stage 1 for all the predictors with the exception of symptomatic lower abdominal pain (OR=1.39[0.19 − 10.39]) and symptomatic vaginal discharge (OR=3.91[0.49 − 31.42]) with low precision and non-statistical significance. The analysis confirms that the FIGO clinical stage is the key factor within the cervical cancer setting when the gynecologists determine the suitability of the surgical treatment therapy for women with confirmed invasive cancer.

In our study,the CR model without the proportional odds assumption was the best fit compared to the CR model with proportional odds. Based on the comparison of models, the work of Guzman-Castillo et al. (2015) compared the continuation ratio model, the adjacent category model, the multinomial model and two other models on the ordinal response of hospital length of stay with patient characteristics as covariates. The ordinal regression model, the CR model and the ACL model violated the proportional odds assumption. Moreover, in their article, Buch et al. (2005) compared the estimated relative risks of the multinomial model, the cumulative ratio model and the continuation ratio model on blood cancer ordinal responses. The authors determined through the goodness-of-fit statistics, the regression diagnostic analysis, small standard errors and smaller 95% confidence intervals that the CR model was the best fit model for the ordinal responses. The work of Dos Santos and Berridge (2000) expounded on the CR model as compared to the ACL model and the baseline category model sharing that the CR model is recognized for being a simple decomposition of a multinomial distribution, its possession of the property of conditional independence between categories and the model's significance levels capability of being affected by a reversal in the order of the categories. In addition, a prior study by Zhou et al. (2008) compared the fit of the baseline category model, the proportional odds model and the adjacent category model in determining the prostate cancer stage and found the baseline category model to have the highest DIC. The authors took the investigation further by comparing the baseline category model to a logistic regression model fitted to dichotomized ordinal responses which demonstrated that the baseline category model was a superior fit. de Jong et al. (2019) recommended that at least 50 multinomial events per variable. The author indicated that the MLR predictive performance gradually improves as the number of multinomial events per variable increases. Our study results show that this could be the possible reason for the MLR model estimated by maximum likelihood being the most unlikely choice among the 3 ordinal regression models.

The Predictive Approach

The predictive approach utilized a larger sample size based on simulation and the results showed that the 5 predictor variables were not statistically significant in predicting the surgical stage outcomes for women with confirmed cervical cancer prior to surgical treatment. The multivariate CR model without proportional odds presented an AIC value of 13754.89 with the lowest deviance of 13730.89 and highest log-likelihood ratio of -6865.44 indicating that it would be the best model to select for the simulated cervical cancer data. In contrast, the predictive approach demonstrated no similarity between the MLR and ACL model which may be attributed to the large sample size utilized to train the models. Clearly, this outcome was in contrast to the inferential approach that showed some similarities between the 2 regression models. In order to test the proportional odds assumption, the CR model and the ACL models were fitted with and without proportional odds. The accuracy of the CR model without proportional odds was 66.78%. However, the goodness-of-fit statistics showed that the CR model without proportional odds gave the lowest deviance of 13730.89 and a low AIC statistic of 13754.89. The predictive approach clearly demonstrated that the 5 predictor variables were not statistically significant in predicting the surgical stage outcomes. Clearly, there is need for collection of additional clinical and laboratory data that can aid the development of predictive algorithms.
The 5 predictor variables gave odds ratios with confidence intervals that did span the null value (OR = 1) demonstrating statistical non-significance. In contrast, the odds ratio results gave small confidence intervals indicative of higher levels of precision (see Table 4.38). The OR for vaginal involvement showed decreased odds of having a surgical stage greater than surgical stage 1 relative to being in surgical stage 1 and decreased odds of having a surgical stage greater than surgical stage 2 relative to being in surgical stage 2. The OR for FIGO clinical stage and parametrial involvement showed increased odds of having a surgical stage greater than surgical stage 1 relative to being in surgical stage 1 and in contrast, decreased odds of having a surgical stage greater than surgical stage 2 relative to being in surgical stage 2.The OR for symptomatic vaginal discharge and symptomatic lower abdominal pain showed decreased odds of having a surgical stage greater than surgical stage 1 relative to being in surgical stage 1 and in contrast, increased odds of having a surgical stage greater than surgical stage 2 relative to being in surgical stage 2.
An accuracy of 66.78%[64.67% $-$ 68.85%] was found for the ACL multivariate model without proportional odds which was slightly lower than the ACL model based on the inferential approach. We obtained similar accuracy for the MLR multivariate model and CR model without proportional odds of 66.78%

[CI:64.67% – 68.86%] and 66.87% [CI:64.67% – 68.85%] respectively.

An AUC provides the overall measure of performance across all possible classification thresholds. The AUC value is considered scale invariant given that it is not an absolute value but a measure of how well predictions are ranked. Moreover, it measures the quality of the model's predictions irrespective of the chosen classification category. The AUC value has a range of 0 – 1 with a value of 0.0 indicative of a model with 100.00% incorrect predictions and a very poor measure of separability. A value of 1.0 is indicative of a model with 100.00% correct predictions and a good measure of separability. We obtained AUC values for the chosen multivariate MLR, CR model with PPO and ACL model without PPO to be approximately 0.5 (see Table 4.46 Therefore, the AUC values obtained indicate that our 3 predictive ordinal regression models of choice have no class separation capabilities.

Gentry et al. (2015) reported that traditional modeling methods were often unsuitable as they required the number of predictors to be smaller than the sample size and the predictors to be independent. However, in his study, Gentry et al. (2015) gave preference to penalized regression models for ordinal responses which introduced bias into the models in exchange for reducing variability. Similar to our study, Gentry et al. (2015) chose the log likelihood and AIC as the measures of the relative quality of the statistical models with addition of the BIC statistic.The author did extract parameter estimates at the point that minimized the AIC value. Through these steps, Gentry et al. (2015) suggested that ordinal regression models should be developed that would allow for the inclusion of non-penalized covariates,covariates or both. The author stated that this can be applied to the CR models or the ACL models. This can be an area for future research utilizing data from the cervical cancer clinics.

5.3 Implications

Based on an inferential and predictive approach, the thesis presented the comparison between 3 different regression models for ordinal data with respect to the best fit model for our cervical cancer data. We found that the CR model without proportional odds yielded better results due to the highest AIC and log likelihood ratio and the lowest residual deviance. In addition, it is clear that based on the inferential approach, the key prognostic factor associated with invasive cervical cancer was the FIGO clinical stage which particularly, had a higher influence on the surgical stage 2 outcomes compared to the lesser surgical stage categories. Of all the 5 independent features selected for classifying the patients into surgical stages, the FIGO clinical stage took precedence. However, the predictive approach showed that neither the FIGO clinical stage was not sufficient to determine the surgical stage outcomes nor the remaining 4 predictors. Thus, we realize that further clinical and laboratory parameters are required to achieve the application of predictive algorithms at the oncology department at CCCDC. Moreover, the predictive approach supports the need to utilize a large sample as it incorporates more precision and leads to efficient odds ratios.

The study was limited by the fact that the cervical cancer data was not created for the purpose of building statistical models thus was not sufficient and probably lacked key predictors for the type of analysis carried out in our study. Thus, our study demonstrates the need of databases with additional variables that could be significant to determining the suitability of surgical treatment such as molecular data, CT / MRI imaging information and HPV-DNA types. In addition, several misclassifications may be attributed to the imbalanced and largely varying category sizes particularly for surgical stages 1 and 2. Moreover, research and data collection for predictive algorithms could introduce practical learning tools for the medical students who undergo medical training at the Moi Teaching and Referral hospital. The data was biased due to the dropping of incomplete records which left a small sample for building the models. Also, data was simulated to test the predictive capabilities of the models and statistical techniques were not utilized to address the imbalanced nature of the data as well as missing data. Although 4 predictors were not found to be key prognostic factors for highly accurate classifications in our models, future research utilizing data structured for developing predictive models in the cervical cancer setting could yield better results that could be integrated into the oncology system. Although actual data was not utilized to validate the models, the fact that the CR model is strictly an ordinal classifier supports what Chen (2012) reported that a strict and validated ordinal classifier can more accurately predict the cancer stages (ordinal scales) compared to non-ordinal classifiers as noted by the multinomial logistic regression model.

References

Abraham, A. G., Strickler, H. D., and D'souza, G. (2013). Invasive cervical cancer risk among hiv-infected women is a function of cd4 count and screening. *Journal of acquired immune deficiency syndromes (1999)*, 63(5):e163.

Ananth, C. V. and Kleinbaum, D. G. (1997). Regression models for ordinal responses: a review of methods and applications. *International journal of epidemiology*, 26(6):1323–1333.

Atashili, J. (2009). *Cervical precancerous lesions in HIV-positive women in Cameroon: prevalence, predictors and potential impact of screening*. PhD thesis, The University of North Carolina at Chapel Hill.

Australia, C. C. (2015). Cancer council australia:understanding cervical cancer a guide for women with cancer, their families and friends. National Publications Working Group initiative. Accessed on September 8th,2017 from http://www.cancervic.org.au/downloads/resources/booklets/Understanding-Cervical-Cancer.pdf.

Averette, H. E., Dudan, R. C., and Ford Jr, J. H. (1972). Exploratory celiotomy for surgical staging of cervical cancer. *American Journal of Obstetrics and Gynecology*, 113(8):1090–1096.

Bhatla, N., Berek, J. S., Fredes, M. C., Denny, L. A., Grenman, S., Karunaratne, K., Kehoe, S. T., Konishi, I., Olawaiye, A. B., Prat, J., et al. (2019). Revised figo staging for carcinoma of the cervix uteri. *International Journal of Gynecology & Obstetrics*, 145(1):129–135.

Biewenga, P., van der Velden, J., Mol, B. W. J., Stalpers, L. J., Schilthuis, M. S., van der Steeg, J. W., Burger, M. P., and Buist, M. R. (2011). Prognostic model for survival in patients with early stage cervical cancer. *Cancer*, 117(4):768–776.

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, pages 1171–1178.

Buch, S. C., Branch, R. A., Jones, K., Arena, V. C., Persad, R., and Romkes, M. (2005). The effect of model choice on the estimates of relative risk for bladder cancer.

Chambuso, R. S., Shadrack, S., Lidenge, S. J., Mwakibete, N., and Medeiros, R. M. (2016). Influence of hiv/aids on cervical cancer: A retrospective study from tanzania. *Journal of global oncology*, 3(1):72–78.

Chen, C.-K. (2012). The classification of cancer stage microarray data. *Computer methods and programs in biomedicine*, 108(3):1070–1077.

Chow, E., Abdolell, M., Panzarella, T., Harris, K., Bezjak, A., Warde, P., and Tannock, I. (2008). Predictive model for survival in patients with advanced cancer. *Journal of Clinical Oncology*, 26(36):5863–5869.

Christensen, R. H. B. (2012). Ordinal—regression models for ordinal data. R package version. Accessed on December 21st,2019 from http://www2.uaem.mx/r-mirror/web/packages/ordinal/.

Coleman, J. S., Cespedes, M. S., Cu-Uvin, S., Kosgei, R. J., Maloba, M., Anderson, J., Wilkin, T., Jaquet, A., Bohlius, J., Anastos, K., et al. (2016). An insight into cervical cancer screening and treatment capacity in sub-saharan africa. *Journal of lower genital tract disease*, 20(1):31.

Cousins, M. S., Shickle, L. M., and Bander, J. A. (2002). An introduction to predictive modeling for disease management risk stratification. *Disease Management*, 5(3):157–167.

de Jong, V. M., Eijkemans, M. J., van Calster, B., Timmerman, D., Moons, K. G., Steyerberg, E. W., and van Smeden, M. (2019). Sample size considerations and predictive performance of multinomial logistic prediction models. *Statistics in medicine*, 38(9):1601–1619.

de Rooij, M. (2019). Logit models for multinomial responses. *Leiden University*.

Denslow, S. A., Rositch, A. F., Firnhaber, C., Ting, J., and Smith, J. S. (2014). Incidence and progression of cervical lesions in women with hiv: a systematic global review. *International journal of STD & AIDS*, 25(3):163–177.

Derks, M., Biewenga, P., van der Velden, J., Kenter, G. G., Stalpers, L. J., and Buist, M. R. (2016). Results of radical surgery in women with stage ib2/iia2 cervical cancer. *Acta obstetricia et gynecologica Scandinavica*, 95(2):166–172.

Dolgun, A. and Saracbasi, O. (2014). Assessing proportionality assumption in the adjacent category logistic regression model. *Statistics and Its Interface*, 7(2):275–295.

Dos Santos, D. M. and Berridge, D. M. (2000). A continuation ratio random effects model for repeated ordinal responses. *Statistics in medicine*, 19(24):3377–3388.

Drolet, M., Boily, M.-C., Greenaway, C., Deeks, S. L., Blanchette, C., Laprise, J.-F., and Brisson, M. (2013). Sociodemographic inequalities in sexual activity and cervical cancer screening: implications for the success of human papillomavirus vaccination. *Cancer Epidemiology and Prevention Biomarkers*, 22(4):641–652.

Fader, A. N. and Rose, P. G. (2007). Role of surgery in ovarian carcinoma. *Journal of clinical oncology*, 25(20):2873–2883.

Freedman, A. N., Seminara, D., Gail, M. H., Hartge, P., Colditz, G. A., Ballard-Barbash, R., and Pfeiffer, R. M. (2005). Cancer risk prediction models: a workshop on development, evaluation, and application. *Journal of the National Cancer Institute*, 97(10):715–723.

Gancarczyk, K. J., Wu, H., McLeod, D. G., Kane, C., Kusuda, L., Lance, R., Herring, J., Foley, J., Baldwin, D., Bishoff, J. T., et al. (2003). Using the percentage of biopsy cores positive for cancer, pretreatment psa, and highest biopsy gleason sum to predict pathologic stage after radical prostatectomy: the center for prostate disease research nomograms. *Urology*, 61(3):589–595.

Gentry, A. E., Jackson-Cook, C. K., Lyon, D. E., and Archer, K. J. (2015). Penalized ordinal regression methods for predicting stage of cancer in high-dimensional covariate spaces. *Cancer informatics*, 14:CIN–S17277.

Girardi, F., Lichtenegger, W., Tamussino, K., and Haas, J. (1989). The importance of parametrial lymph nodes in the treatment of cervical cancer. *Gynecologic oncology*, 34(2):206–211.

Gold, M. A., Tian, C., Whitney, C. W., Rose, P. G., and Lanciano, R. (2008). Surgical versus radiographic determination of para-aortic lymph node metastases before chemoradiation for locally advanced cervical carcinoma: a gynecologic oncology group study. *Cancer*, 112(9):1954–1963.

Grace-Martin, K. (2019). Logistic regression models for multinomial and ordinal variables. The Analysis Factor. Accessed on August 30th,2019 from https://www.theanalysisfactor.com/logistic-regression-models-for-multinomial-and-ordinal-variables/.

Guzman-Castillo, M., Brailsford, S., Luke, M., and Smith, H. (2015). A tutorial on selecting and interpreting predictive models for ordinal health-related outcomes. *Health Services and Outcomes Research Methodology*, 15(3-4):223–240.

Heaps, J. M. and Berek, J. S. (1990). Surgical staging of cervical cancer. *Clinical obstetrics and gynecology*, 33(4):852–862.

Hsu, H.-C., Tai, Y.-J., Chen, Y.-L., Chiang, Y.-C., Chen, C.-A., and Cheng, W.-F. (2018). Factors predicting parametrial invasion in patients with early-stage cervical carcinomas. *PloS one*, 13(10):e0204950.

Huang, Z., Mayr, N. A., Yuh, W. T., Lo, S. S., Montebello, J. F., Grecula, J. C., Lu, L., Li, K., Zhang, H., Gupta, N., et al. (2010). Predicting outcomes in cervical cancer: a kinetic model of tumor regression during radiation therapy. *Cancer research*, pages 0008–5472.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Joniau, S., Hsu, C.-Y., Lerut, E., Van Baelen, A., Haustermans, K., Roskams, T., Oyen, R., and Van Poppel, H. (2007). A pretreatment table for the prediction of final histopathology after radical prostatectomy in clinical unilateral t3a prostate cancer. *European urology*, 51(2):388–396.

Kamura, T., Tsukamoto, N., Tsuruchi, N., Saito, T., Matsuyama, T., Akazawa, K., and Nakano, H. (1992). Multivariate analysis of the histopathologic prognostic factors of cervical cancer in patients undergoing radical hysterectomy. *Cancer*, 69(1):181–186.

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., and Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology journal*, 13:8–17.

Lagasse, L. D., Creasman, W. T., Shingleton, H. M., Ford, J. H., and Blessing, J. A. (1980). Results and complications of operative staging in cervical cancer: experience of the gynecologic oncology group. *Gynecologic oncology*, 9(1):90–98.

Lans, D. (2019). Illustrating predictive models with the roc curve. Medium. Accessed on August 24th,2020 from http://towardsdatascience.com/illustrating-predictive-models-with-the-roc-curve-67e7b3aa8914.

Lee, C.-H., Peng, C.-Y., Li, R.-N., Chen, Y.-C., Tsai, H.-T., Hung, Y.-H., Chan, T.-F., Huang, H.-L., Lai, T.-C., and Wu, M.-T. (2015). Risk evaluation for the development of cervical intraepithelial neoplasia: Development and validation of risk-scoring schemes. *International journal of cancer*, 136(2):340–349.

Liu, X. (2010). Ordinal regression analysis: Fitting the continuation ratio model to educational data using stata.

Marnitz, S., Köhler, C., Roth, C., Füller, J., Hinkelbein, W., and Schneider, A. (2005). Is there a benefit of pretreatment laparoscopic transperitoneal surgical staging in patients with advanced cervical cancer? *Gynecologic oncology*, 99(3):536–544.

Matsuo, K., Mabuchi, S., Okazawa, M., Kawano, M., Kuroda, H., Kamiura, S., and Kimura, T. (2015). Clinical implication of surgically treated early-stage cervical cancer with multiple high-risk factors. *Journal of gynecologic oncology*, 26(1):3–11.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42(2):109–127.

Meisner, A., Parikh, C. R., and Kerr, K. F. (2018). Using ordinal outcomes to construct and select biomarker combinations for single-level prediction. *Diagnostic and prognostic research*, 2(1):8.

Meloun, M. and Militky, J. (2011). *Statistical data analysis: A practical guide*. Woodhead Publishing, Limited.

Nordqvist, C. (2017). Cervical cancer: Causes, symptoms, and treatments. Healthline Media UK Ltd. Accessed on August 28th,2017 from http://www.medicalnewstoday.com/articles/159821.php.

Norton, E. C. and Dowd, B. E. (2018). Log odds and the interpretation of logit models. *Health services research*, 53(2):859–878.

Partin, A. W., Kattan, M. W., Subong, E. N., Walsh, P. C., Wojno, K. J., Oesterling, J. E., Scardino, P. T., and Pearson, J. (1997). Combination of prostate-specific antigen, clinical stage, and gleason score to predict pathological stage of localized prostate cancer: a multi-institutional update. *Jama*, 277(18):1445–1451.

Pecorelli, S. and Odicino, F. (2003). Cervical cancer staging. *The Cancer Journal*, 9(5):390–394.

Petry, K. U. (2014). Hpv and cervical cancer. *Scandinavian Journal of Clinical and Laboratory Investigation*, 74(sup244):59–62.

Portugués, E. G. (2020). Notes for predictive modeling. Carlos III University of Madrid. Accessed on February 17th,2020 from https://bookdown.org/egarpor/PM-UC3M/.

Powers, C. A., Meyer, C. M., Roebuck, M. C., and Vaziri, B. (2005). Predictive modeling of total healthcare costs using pharmacy claims data: a comparison of alternative econometric cost modeling techniques. *Medical care*, pages 1065–1072.

Pradjatmo, H. (2015). Basic principles and diagnostic of colposcopy. 47:102.

Prasad, T., Thulkar, S., Hari, S., Sharma, D., and Kumar, S. (2014). Role of computed tomography (ct) scan in staging of cervical carcinoma. *The Indian journal of medical research*, 139(5):714.

Qin, Y., Peng, Z., Lou, J., Liu, H., Deng, F., and Zheng, Y. (2009). Discrepancies between clinical staging and pathological findings of operable cervical carcinoma with stage ib–iib: A retrospective analysis of 818 patients. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 49(5):542–544.

Ripley, B., Venables, W., and Ripley, M. B. (2016). Package 'nnet'. *R package version*, 7:3–12.

Rodríguez, G. (2019a). Generalized linear models. Princeton University. Accessed on December 27th,2019 from https://data.princeton.edu/wws509/r/c6s2.

Rodríguez, G. (2019b). Generalized linear modelslogistic regression models for multinomial and ordinal variables. Princeton University. Accessed on August 30th,2019 from https://data.princeton.edu/wws509/notes/c6s2.

Rosenfeld, B. A. and Breslow, M. (2008). Using predictive models to continuously update a treatment plan for a patient in a health care location. US Patent 7,395,216.

Rothberg, M. (2018). Cleveland clinic develops first overall risk prediction model for cervical cancer. clevelandclinic.org. Accessed on December 12th,2018 from https://consultqd.clevelandclinic.org/cleveland-clinic-develops-first-overall-risk-prediction-model-for-cervical-c

Rutledge, T. L., Kamelle, S. A., Tillmanns, T. D., Gould, N. S., Wright, J. D., Cohn, D. E., Herzog, T. J., Rader, J. S., Gold, M. A., Johnson, G. A., et al. (2004). A comparison of stages ib1 and ib2 cervical cancers treated with radical hysterectomy. is size the real difference? *Gynecologic oncology*, 95(1):70–76.

Sangwa-Lugoma, G., Ramanakumar, A. V., Mahmud, S., Liaras, J., Kayembe, P. K., Tozin, R. R., Lorincz, A., and Franco, E. L. (2011). Prevalence and determinants of high-risk human papillomavirus infection in women from a sub-saharan african community. *Sexually transmitted diseases*, 38(4):308–315.

Sayed, S. (2019). Machine learning is the future of cancer prediction. A Medium Corporation. Accessed on July 28th,2019 from https://towardsdatascience.com/machine-learning-is-the-future-of-cancer-prediction-e4d28e7e6dfa.

Seel, R. T., Steyerberg, E. W., Malec, J. F., Sherer, M., and Macciocchi, S. N. (2012). Developing and evaluating prediction models in rehabilitation populations. *Archives of physical medicine and rehabilitation*, 93(8):S138–S153.

Sharma, M. (2019). Multinomial logistic regression using r. Data Science Beginners. Accessed on October 26th,2019 from https://datasciencebeginners.com/2018/12/20/multinomial-logistic-regression-using-r/.

Sonia, M., Rodolfo, R., Harmon, S. G., Mabeya, H., and Callens, S. (2017). Public health approach to prevent cervical cancer in hiv-infected women in kenya: Issues to consider in the design of prevention programs. *Gynecologic oncology reports*.

Starkweather, J. and Moske, A. K. (2011). Multinomial logistic regression. *Consulted page at September 10th: http://www. unt. edu/rss/class/Jon/Benchmarks/MLR_JDS_Aug2011. pdf*, 29:2825–2830.

Sun, J., Zhang, Y.-N., Sun, X., Feng, S., and Yan, M. (2011). Prediction model of pelvic lymph node metastasis in early stage cervical cancer and its clinical value. *Minerva chirurgica*, 66(6):537–545.

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian academy of child and adolescent psychiatry*, 19(3):227.

van Nagell Jr, J. R., Roddick Jr, J., and Lowin, D. (1971). The staging of cervical cancer: inevitable discrepancies between clinical staging and pathologic findings. *American journal of obstetrics and gynecology*, 110(7):973–978.

Vickers, A. J. (2011). Prediction models in cancer care. *CA: a cancer journal for clinicians*, 61(5):315–326.

Vinci, L. D. (2014). Clinical presentation, stages and treatment of cervical cancer. eurocytology. Accessed on January 26th,2019 from https://www.eurocytology.eu/en/course/1022.

Wan Kai, P. (2008). Continuation-ratio model for categorical data: A gibbs sampling approach. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1.

Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology*, 40(1):7–20.

Wilson, K. L., Cowart, C. J., Rosen, B. L., Pulczinski, J. C., Solari, K. D., Ory, M. G., and Smith, M. L. (2018). Characteristics associated with hpv diagnosis and perceived risk for cervical cancer among unmarried, sexually active college women. *Journal of Cancer Education*, 33(2):404–416.

Winship, C. and Mare, R. D. (1984). Regression models with ordinal variables. *American sociological review*, pages 512–525.

Yuri, P., Rochadi, S., and Danarto, R. (2016). A device for predicting prostate cancer risk: A logistic regression. *Journal of Prostate Cancer*, pages 1–5.

Zhou, H., Lawson, A. B., Hebert, J. R., Slate, E. H., and Hill, E. G. (2008). A bayesian hierarchical modeling approach for studying the factors affecting the stage at diagnosis of prostate cancer. *Statistics in medicine*, 27(9):1468–1489.