ANALYZING CERVICAL CANCER CASES AT GWERU DISTRICT HOSPITAL IN ZIMBABWE: USING ARTIFICIAL NEURAL NETWORKS

Dr. Smartson. P. NYONI[1], Thabani NYONI[2]

[1] *Medical Doctor, ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe*
[2] *MSc Economics Scholar, Department of Economics, University of Zimbabwe, Harare, Zimbabwe*

ABSTRACT

In this research paper, ANN models are applied to forecast cervical cancer cases seen and managed at Gweru District Hospital (GDH), Gweru, Zimbabwe. The study used a data set covering the period January 2010 – December 2019. The out-of-sample period covers the period January 2020 – December 2021. The applied neural network model is stable and thus suitable for modeling and forecasting cervical cancer cases at GDH. The results of the study generally indicate that monthly cervical cancer cases are likely to be 0, 1 or 2 in the GDH catchment area over the period January 2020 – December 2021. In order to reduce cervical cancer incidence to the barest minimum of 0 cases per month, the paper recommends regular mass cervical cancer screening, amongst other recommendations.

Keyword: - *ANN, Cervical Cancer, Forecasting*

I. INTRODUCTION

Cervical cancer is a reality and will remain with us, as a high impact risk to public health status of females (Mhandu *et al*., 2017) and is actually the second most common cancer in women worldwide (WHO, 2010) and the first in under-developed and developing countries (Bray *et al*., 2018) and apparently remains an important reason of deaths worldwide (Hejmadi, 2009; ACS, 2011; Torre *et al*., 2015; Fernandes *et al*., 2015; Langat *et al*., 2017; Fernandes *et al*., 2018; Benazir & Nagarajan, 2018; Al-Wesabi *et al*., 2018; Alam *et al*., 2019) despite the fact that it can be prevented and cured by removing affected tissues in early stages (Kauffman *et al*., 2013; CDC, 2013; Plissitti & Nikou, 2013; Fernandes *et al*., 2015; Xu *et al*., 2016; Fernandes *et al*., 2017b; Fernandes *et al*., 2018). The expected cancer incidences in the world will reach up to 22 million by 2030 (Hejmadi, 2009; Kamil & Kamil, 2015). There are 4.7 million women aged 15 years and above who are vulnerable to cervical cancer in Zimbabwe (ICO, 2015). Cancer of the cervix is the most frequent diagnosed cancer among females in Zimbabwe (Bassett *et al*., 1995). In fact, cervical cancer is the leading malignancy in females in Zimbabwe (Zimbabwe Cancer Association, 2009; Ministry of Health & Child Care, 2013). In Zimbabwe, 2270 new cervical cancer cases are diagnosed annually with 1451 cervical cancer deaths occurring annually (HVP Centre, 2012), making it the most deadly cancer among Zimbabwean women (Panganai & Gono, 2017).

Millions of early deaths among women are due to lung and breast cancer but cervical cancer is the most dangerous because it is only diagnosed in females. A woman's reproductive system consists of cervix, uterus, vagina and the ovaries. The cervix is the opening to the uterus from the vagina where cervical cancer occurs (Subramanian *et al*., 2016). Sexually transmitted Human Papillomavirus (HPV) is the major cause of cervical cancer (Bassett *et al*., 1995; WHO, 2013; Petry, 2014; Ronco *et al*., 2014). Zimbabwean population proportion age group is between 16-40 years, which happens to be the sexually active age group at high risk of this cancer (Chokunonga *et al*., 2013). Cervical cancer is also associated with risk factors such as socio-economic context, health behavior and smoking (Gonzaga *et al*., 2013).

Cervical cancer develops in the cervix of the women (Healthline, 2020). The development of cervical cancer is usually slow and preceded by abnormalities in the cervix (dysplasia). However, the absence of early stage symptoms might cause carelessness in prevention. Moreso, in developing countries, there is a lack of resources and patients usually have poor adherence to routine screening due to low problem awareness (Fernandes *et al*., 2018). The screening method which is manually done for cervical cells collection in the cervix region of the uterus is the Pap smear test. Cervical cancer can also be usually found early by having regular screening with this Pap test. Being alert to any signs and symptoms of cervical cancer can play a pivotal role in avoiding unnecessary delays in diagnosis (Healthline, 2020). Early detection dramatically improves the chances of successful treatment and prevents any early cervical cell changes from becoming cancerous. In fact, it is better to have a regular screening of cervical cancer because women with early cervical cancer and pre-cancers usually have no symptoms, and symptoms will arise only after the invasive growth of cancerous cells. But for most of the cases, it throw symptoms only in the advanced stages (Benazir & Nagarajan, 2018).

Against this background, the precise forecasting of cancer trends and the estimation of future incidence rates allow better resource allocation, which is invaluable from a public health standpoint (Sathian *et al*., 2013). This study will consolidate Gweru District Hospital (GDH)'s efforts in cervical cancer prevention, screening, diagnosis, treatment and care as well as investment needed to effectively deliver cervical cancer prevention programmes. In fact, this is the first cervical cancer control model to be done for the GDH catchment area and is expected to steer up a scholarly debate in light of the fight against cervical cancer in Zimbabwe.

II. LITERATURE REVIEW

Sathian *et al* (2013) estimated the numbers and trends in cervix cancer cases visiting the Radiotherapy Department at Manipal Teaching Hospital, Pokhara, Nepal; statistical modelling from retrospective data was applied. Their results indicate that there will be a significant increase in cervical cancer cases in Western Nepal in the near future. In China, Du *et al*. (2015) analyzed cervical cancer mortality trends over the period 1991-2013. Trend-surface analysis was used to analyze the geographical distribution of mortality. Curve estimation, time series, gray modeling and jointpoint regression were performed to predict and forecast mortality trends. The study showed that the incidence rate and the mortality rate are increasing from 1991 to 2013, and the predictions show this will continue in the future. Langat *et al*. (2017) modeled and forecasted cancer cases in Kenya using the Box-Jenkins ARIMA model and found out that the optimal model for forecasting cancer cases in Kenya is the ARIMA (2, 1, 0) and the model basically predicted an increase in cancer cases in Kenya in the near future.

In Venezuela, Fernandes *et al*. (2018) applied supervised deep learning embeddings for the prediction of cervical cancer diagnosis. The authors presented a computationally automated strategy for predicting the outcome of the patient biopsy, given risk patterns from individual medical records. The study proposed a machine learning technique that allowed a joint and fully supervised optimization of dimensionality reduction and classification models. The study found out that the AUC outperforms previously developed methods, such as denoising autoencoders. In India, Benazir & Nagarajan (2018) used data mining techniques to predict cervical cancer and found out that data mining techniques such as Feature Selection and Classification Techniques are good at predicting cervical cancer in India. In the United States, Al-Wesabi *et al*. (2018) presented diverse classification techniques and shows the advantage of feature selection approaches to the best predicting of cervical cancer disease. Their study showed that age, first sexual intercourse, number of pregnancies, smokes, hormonal contraceptives and STDs are the main predictive features with high accuracy with 97.5%. Decision Tree classifiers is shown to be advantageous in handling classification assignment with excellent performance.

Alam *et al*. (2019) analyzed cervical cancer data in Pakistan using data mining techniques such as Boosted Decision Tree, Decision Forest and Decision Jungle algorithms. The accucary of the predictive models were checked using the Area Under Receiver Operating Characteristic (AUROC) curve. 10-fold cross-validation method was utilized to authenticate the results and Boosted Decision Tree produced best results. Boosted Decision Tree provided very high prediction with 0.978 on AUROC curve while Hinslemann screening method was used. The results obtained by other classifiers were significantly worse than Boosted Decision Tree.

Literature review shows that only two studies model and forecast cancer cases, that is, Sathian *et al*. (2013) and Langat *et al*. (2017). This study is closely related to Sathian *et al*. (2013) even though we differ in terms of

methodology. While Sathian *et al*. (2013) used various time series forecasting methods such as exponential growth models, this paper will apply Artificial Neural Networks (ANNs) which are well known for their better forecast performance.

III. METHODOLOGY

There is an increasingly growing interest in the domain of data mining techniques, especially, Artificial Neural Networks (ANNs) for analyzing various cervical cancer data sets (Du *et al*., 2015; Fernandes *et al*., 2018; Benazir & Nagarajan, 2018; Al-Wesabi *et al*., 2018). This study applies the multi-layer perceptron neural network type of the Artificial Neural Network technique in order to analyze cervical cancer cases seen and recorded at GDH.

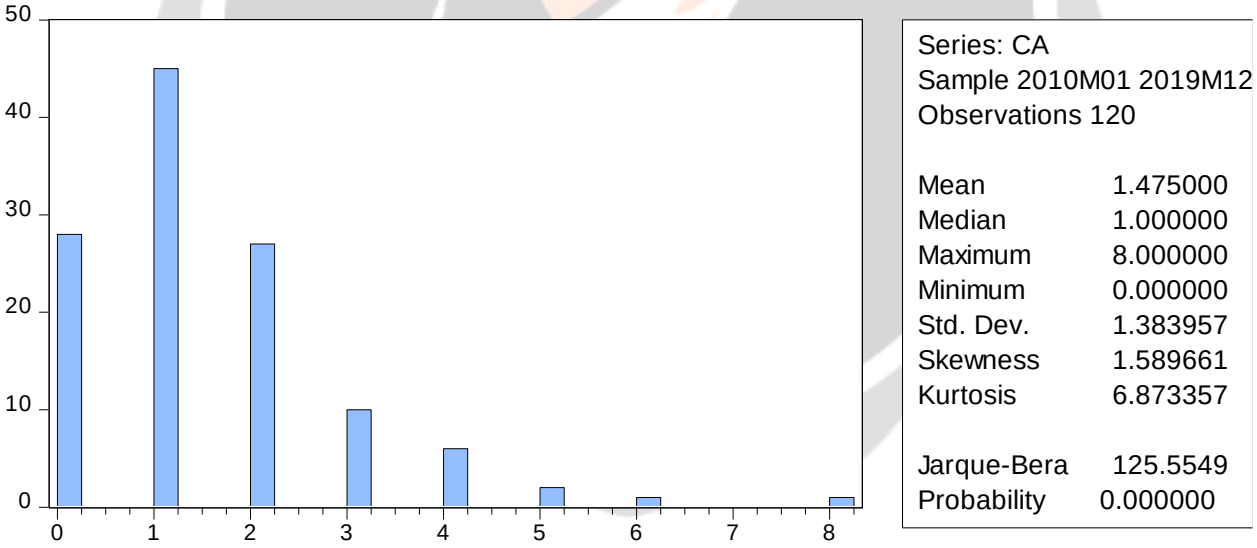
Data Issues

This study is based on newly diagnosed monthly cervical cancer cases (reffered to as CA series in this study) in women aged 45 years and above at GDH. The data covers the period January 2010 to December 2019 while the out-of-sample forecast covers the period January 2020 to December 2021. All the data employed in this paper was gathered from GDH Health Information Department.

IV. FINDINGS OF THE STUDY

Descriptive Statistics

Figure 1: Descriptive statistics



The average number of cervical cancer cases over the study period is approximately 1 case per month while the maximum number of cervical cancer cases is 8. While these figures are not really shocking, it is desirable, one day to observe 0 maximum not minimum cervical cancer cases per month. This possible because cervical cancer is preventable and curable if detected and treated urgently.

ANN Model Summary for Cervical Cancer Cases at GDH

Table 1: ANN model summary

Variable	CA
Observations	108 (After Adjusting Endpoints)
Neural Network Architecture:	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Function	Hyperbolic Tangent Function
Back Propagation Learning:	
Learning Rate	0.005
Momentum	0.05
Criteria:	
Error	0.128607
MSE	0.326714
MAE	0.455853

Table 1 above is a summary of the estimated ANN model.

Residual Analysis for the ANN model

Figure 2: Residual analysis

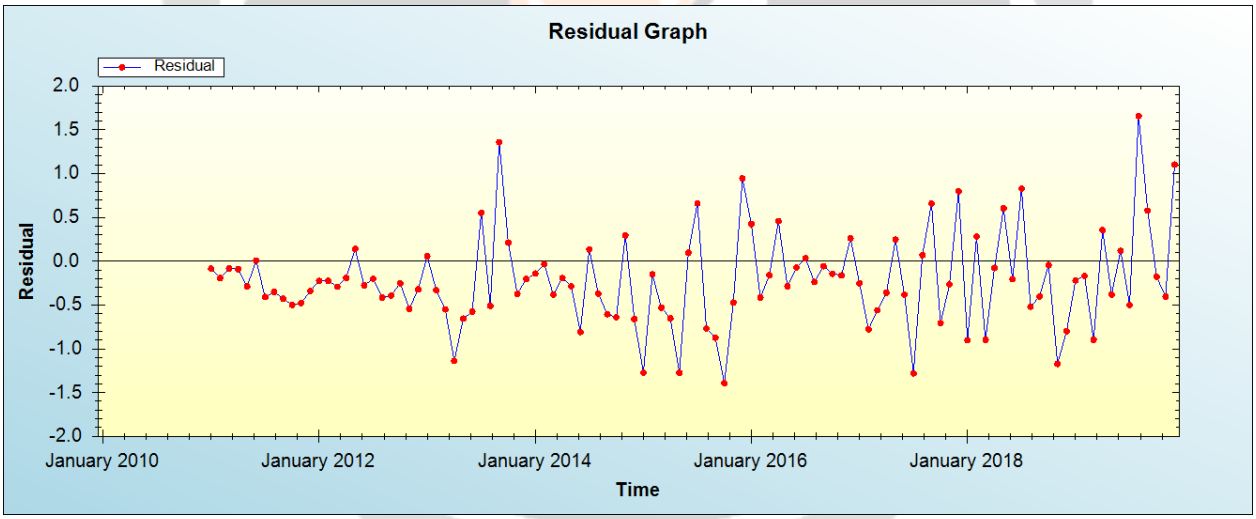


Figure 2 is the residual analysis of the applied ANN model. The residuals are generally revolving around 0, indicating that the applied ANN model is stable and adequate.

In-sample Forecast for CA

Figure 3: In-sample forecast for the CA series

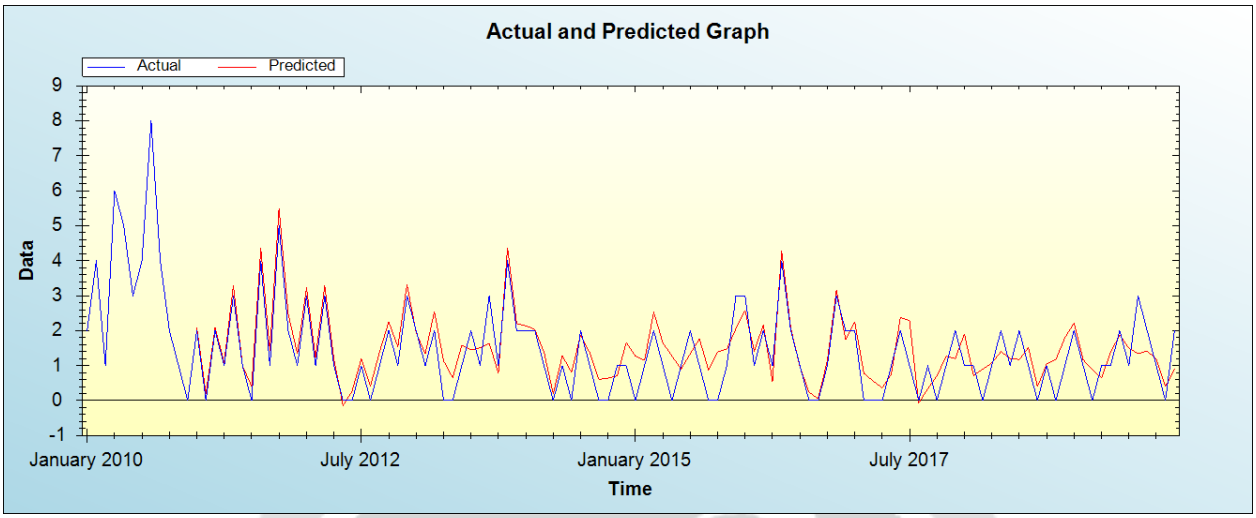


Figure 3 is the in-sample forecast of the CA series.

Out-of-Sample Forecast for CA: Actual and Forecasted Graph

Figure 4: Out-of-sample forecast for CA: actual and forecasted graph

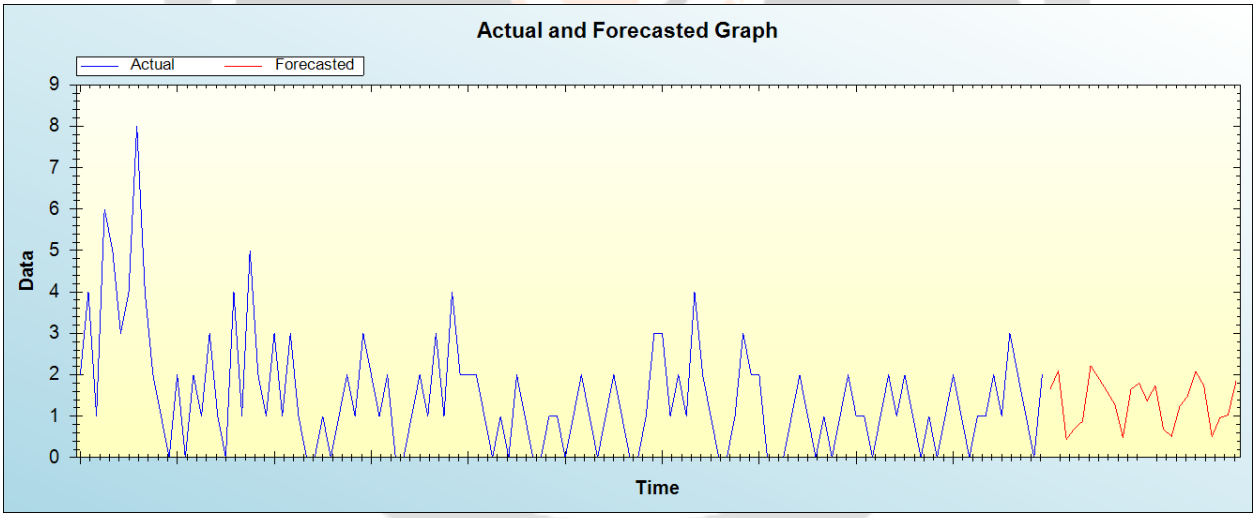


Figure 4 shows the out-of-sample forecasts of the series under consideration. Table 2 below shows the same.

Out-of-Sample Forecast for CA: Forecasts only

Table 2: Tabulated out-of-sample forecasts

Month/Year	Predicted CA
January 2020	1.6582
February 2020	2.0922
March 2020	0.4383

April 2020	0.7046
May 2020	0.8748
June 2020	2.2172
July 2020	1.9224
August 2020	1.6042
September 2020	1.2826
October 2020	0.4892
November 2020	1.6553
December 2020	1.8009
January 2021	1.3693
February 2021	1.7412
March 2021	0.6916
April 2021	0.5108
May 2021	1.2320
June 2021	1.4881
July 2021	2.0807
August 2021	1.7416
September 2021	0.5099
October 2021	0.9615
November 2021	1.0316
December 2021	1.8488

Figure 5: Graphical presentation of out-of-sample forecasts

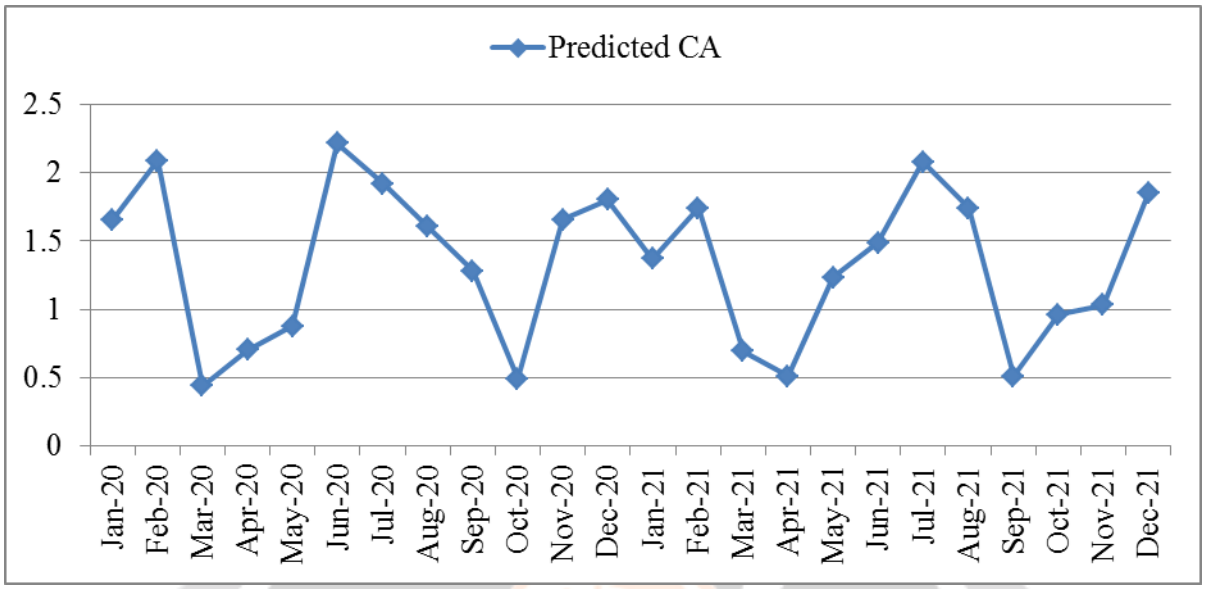


Figure 5 shows the graphical presentation of the predicted cervical cancer cases over the period January 2020 – December 2021. The trend is generally declining over time, although quite slowly. Over the out-of-sample period, a maximum of 2 cervical cancer cases is likely to be seen at GDH. What is impressive is that the numbers are quite minimal: this shows that it is possible to win the war against cervical cancer not only in the GDH catchment area but also in Zimbabwe. In order to guarantee a win against cervical cancer, the following recommendations should be taken seriously:

Recommendations

i. GDH is encouraged to do regular mass cervical screening in its catchment area.

ii. There is need for continued HPV vaccination.

iii. Early treatment of pre-cancer lesions.

iv. The government of Zimbabwe, through the GDH management team, should avail resources for effective diagnosis and treatment of cervical cancer.

v. It is important for GDH to hold educational programs in its catchment area in order to sensitize people about cervical cancer, because it is real. In this regard, there is need for GDH to encourage people to have only one faithful sexual partner as opposed to multiple sexual partners. Furthermore, correct and consistent use of condoms during sexual encounters is greatly encouraged.

vi. GDH medical officers and nurses should see to it that patients do not engage in "prolonged" use oral contraceptive pills.

V. CONCLUSION

Cervical cancer is a deadly disease that affects females of all walks of life regardless of racial, socio-economic status or religious and cultural beliefs (Mhandu *et al*., 2017). Therefore, the ability of forecasting cervical cancer incidence accurately is a major milestone in the control and management of the disease. This study examined monthly cervical cancer cases seen and managed at GDH over the period 2010 – 2019 using ANN models. The results of this study will go a long way in facilitating optimal distribution of resources, enabling the adoption of proper control interventions tailor-made for the GDH catchment area. This is also expected to reduce the number of new cervical

cancer cases and cervical cancer-attributable deaths not only in the GDH catchment area but also in the whole country in the long run.

REFERENCES

[1] Alam, T. M., *et al*. (2019). Cervical Cancer Prediction Through Different Screening Methods Using Data Mining, *International Journal of Advanced Computer Science and Applications*, 10 (2): 388 – 396.

[2] Al-Wesabi, Y. M. S., Choudhury, A., & Won, D. (2018). Classification of Cervical Cancer Dataset, *Proceedings of the 2018 Annual Conference*, pp: 1 – 6.

[3] American Cancer Society (ACS) (2011). Cancer in Africa, *ACS*, New York.

[4] Bassett, M. T., *et al*. (1995). Cancer in the African Population of Harare, Zimbabwe, 1990-1992, *International Journal of Cancer*, 63 (1): 29 – 36.

[5] Benazir, B., & Nagarajan, A. (2018). An Expert System for Predicting the Cervical Cancer Using Data Mining Techniques, *International Journal of Pure and Applied Mathematics*, 118 (20): 1971 – 1987.

[6] Bray, F., *et al*. (2018). Global Cancer Statistics, *GLOBOCAN*, New York.

[7] Centers for Disease Control (CDC) (2013). Cervical Cancer Screening Among Women Aged 18 – 30 Years, United States, 2000 – 2010, *Morbidity and Mortality Weekly Report*, 61 (51): 10 – 38.

[8] Chokunonga, E., *et al*. (2013). Trends in the Incidence of Cancer in the Black Population of Harare, Zimbabwe, 1991-2010, *International Journal of Cancer*, 133 (3): 721 – 730.

[9] Du, P., *et al*. (2015). Cervical Cancer Mortality Trends in China, 1991 – 2013, and Predictions for the Future, *Asia Pacific Journal of Cancer Prevention*, 16 (15): 6391 – 6396.

[10] Fernandes, K., Cardoso, J. S., & Fernandes, J. (2015). Temporal Segmentation of Digital Colposcopies, *Springer*, pp: 262 – 271.

[11] Fernandes, K., Cardoso, J. S., & Fernandes, J. (2017b). Transfer Learning With Partial Observability Applied to Cervical Cancer Screening, *Springer*, 243 – 250.

[12] Fernandes, K., *et al*. (2018). Supervised Deep Learning Embeddings for the Prediction of Cervical Cancer Diagnosis, *Computer Science*, 4 (154): 1 – 20.

[13] Gonzaga, C. M. R., *et al*. (2013). Cervical Cancer Mortality Trends in Brazil: 1980-2009, *Cad Saude Publica*, 29: 599 – 608.

[14] Healthline (2020). Pap-smear, *Healthline*, London.

[15] Hejmadi, M. (2009). Introduction to Cancer Biology, Bookboon.

[16] HPV Centre (2012). Human Papillomavirus and Related Diseases, *HPV Centre*, New York.

[17] ICO (2015). HPV and Related Cancers, *ICO*, New York.

[18] Kauffman, R. P., *et al*. (2013). Current Recommendations for Cervical Cancer Screening: Do they Render the Annual Pelvic Examination Obsolete? *Medical Principles and Practice*, 22 (4): 313 – 322.

[19] Langat, A., Orwa, G., & Koima, J. (2017). Cancer Cases in Kenya; Forecasting Incidents Using Box & Jenkins ARIMA Model, *Biomedical Statistics and Informatics*, 2 (2): 37 – 48.

[20] Mhandu, T., *et al*. (2017). Conceptualization of Cervical Cancer Causation and Risk Factors: A Qualitative Investigation of Women in the City of Harare, Zimbabwe, *Scholars Journal of Applied Medical Sciences*, 5 (7): 2722 – 2731.

[21] Ministry of Health & Child Care (2013). Maternal and Child Care, *Government of Zimbabwe*, Harare.

[22] Panganai, T., & Gono, C. (2017). Visual Inspection With Acetic Acid Screening for Cervical Cancer: Perceptions of Zimbabwean Women: A Case of Kwekwe Hospital, *International Journal of Sport, Exercise and Health Research*, 1 (1): 12 – 17.

[23] Petry, K. U. (2014). HPV and Cervical Cancer, *Scandinavian Journal of Clinical and Laboratory Investigation*, 74: 59 – 62.

[24] Plissiti, M. E., & Nikou, C. (2013). A Review of Automated Techniques for Cervical Cell Image Analysis and Classification, *Springer*, pp: 1 – 18.

[25] Ronco, G., *et al*. (2014). Efficacy of HPV-based Screening for Prevention of Invasive Cancer: Follow-up of Four European Randomized Controlled Trials, *The Lancet*, 383: 524 – 532.

[26] Sathian, B., *et al*. (2013). Statistical Modeling and Forecasting of Cervix Cancer Cases in Radiation Oncology Treatment: A Hospital Based Study From Western Nepal, *Asian Pacific Journal of Cancer Prevention*, 14: 2097 – 2100.

[27] Subramanian, S., *et al*. (2016). Clinical Trial to Implementation: Cost and Effectiveness Considerations for Scaling up Cervical Cancer Screening in Low and Middle-income Countries, *Journal of Cancer Policy*, 7: 4 – 11.

[28] Torre, L. A., *et al*. (2015). Global Cancer Statistics, *Cancer Journal*, 65: 87 – 108.

[29] WHO (2010). Human Papillomavirus Infection and Cervical Cancer, *WHO*, Geneva.

[30] WHO (2013). Comprehensive Cervical Cancer Prevention and Control: A Healthier Future for Girls and Women, *WHO*, Geneva.

[31] Xu, T., *et al*. (2016). Multimodal Deep Learning for Cervical Dysplasia Diagnosis, *Springer*, 115 – 123.

[32] Zimbabwe Cancer Association (2009). Cancer Report, *Zimbabwe Cancer Association*, Harare.