European Scholar Journal (ESJ)
Available Online at: https://www.scholarzest.com
Vol. 1 No. 1, September 2020,
ISSN: 2660-5562

USING ARTIFICIAL NEURAL NETWORKS FOR PREDICTING NEW ESOPHAGEAL CANCER CASES AT GWERU PROVINCIAL HOSPITAL IN ZIMBABWE

Dr. Smartson. P. NYONI
ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe

Mr. Thabani NYONI
Department of Economics, University of Zimbabwe, Harare, Zimbabwe

Article history:

Received August, 8th 2020
Accepted: September 4th 2020
Published: September 30th 2020

Abstract:

Cancer is a disease that affects large numbers of people from all walks of life. Diagnosis of cancer induces fear both in the individual and in families, and is frequently viewed as death sentence (Ministry of Health & Child Care, 2014). Its prevention, diagnosis and treatment poses a myriad of challenges particulary in developing countries such as Zimbabwe. This study employed monthly time series data on Esophageal Cancer (EC) cases at Gweru Provincial Hospital (GPH) from Janaury 2010 to December 2019, in order to predict EC cases over the period January 2020 to December 2021. The ANN (12, 12, 1) model was applied. Residual analysis of this model indicated that it was indeed stable and thus suitable for predicting EC cases at GPH over the out-of-sample period. The results of the study indicate that EC cases will generally rise over the out-of-sample period; characterised by possible seasonal peaks in the month of November each year. Amongst other recommendations, the study suggested the need for prioritizing proper diets and healthy lifestyles as a preventive measure against EC.

Keywords: Series data on Esophageal Cancer (EC), healthy lifestyles.

1.0 INTRODUCTION

Esophageal cancer (EC) is the 6th most common cause of mortality among all cancer worldwide, and the 7th most common cancer worldwide, with approximately 572034 new cases (3.2% of all cancers) and 508585 cancer deaths (5.3% of all cancer deaths) in 2018 (GLOBOCAN, 2018). The incidence of EC in the world is relatively high and on the increase year by year (Hou *et al.*, 2017). In males, it is the 7th most common cancer and men are up to 4 times more at risk of developing EC than women (Tiasto *et al.*, 2018). The prevalence of EC varies geographically and among age and race groups, with the highest risk groups being individuals over 70 years old, and black males (Cancer Research, 2017). More than 80% of EC cases occur in developing countries: the prevalence remains highest in Asia and Africa, while occurrences in North America and Europe are on an exponential rise (Zhang, 2013). Currently, over 5000 new cancer cases are diagnosed (all types) in Zimbabwe annually. Of those who do report, the majority are already at an advanced stage of disease, due to limited access to screening services. The current cancer treatment and palliation services are unable to meet the existing demand in Zimbabwe (Ministry of Health & Child Care, 2014).

The dominant histological subtype of EC is esophageal squamous cell carcinoma (ESCC), followed by adenocarcinoma (AC) (Arnold *et al.*, 2012). ESCC basically arises from cigarette smoking and alcohol consumption (Tiasto *et al.*, 2018) as well as genetic variations of low-activity ethanol-metabolizing enzymes and human papillomavirus infection (Watanabe, 2015) and is the most prevalent type of EC worldwide, responsible fro more than 80% of all EC cases (Tiasto *et al.*, 2018). Other causes of ESCC include consumption of hot beverages, nutritional deficiencies and limited intake of fruits and vegetables. The prognosis for ESCC is poor, with a 5-year survival rate of 19% and only 0.9% for advanced ESCC (Testa *et al.*, 2017). AC arises from the metaplastic Barret's esophagus (BE) in the context of chromic inflammation secondary to exposure to acid and bile (Tiasto *et al.*, 2018). The main risk factors for AC are gastro-esophageal reflux disease and obesity, both leading to the only described precursor lesion for this cancer, namely Barret's esophagus (Wang, 2015). While it is rare in Asia and Africa, AC is highly prevalent in Western countries (Palethorpe *et al.*, 2017), partially due to the concurrent epidemic of obesity (Rubenstein & Shaheen, 2015). In Zimbabwe, physical inactivity and obesity are the main causes of EC (Ministry of Health & Child Care, 2014). Patients with EC may have unspecific symptoms like tiredness, nausea, weight loss and so on., at an early stage, which makes it difficult to diagnose (Hou *et al.*, 2017).

Management of EC includes surgery, radiation, chemotherapy and laser therapy (Kachala, 2010; Middleton *et al.*, 2018). There is still no effective treatment for this deadly disease and the patient's survival remains very poor (Hou *et*

al., 2017).The large number of people living with HIV results in an even higher number of people who will develop cancer in Zimbabwe (Ministry of Health & Child Care, 2014). Meeting this increased demand and ensuring sufficient quality of services will require accurate forecasting models for analyzing cancer trends and forecasting its future evolution.

1.1 OBJECTIVES OF THE STUDY

i. To assess new EC cases in adults aged 45 years and above at GPH over the period January 2010 to December 2019.
ii. To predict EC cases for GPH over the period January 2020 to December 2021.
iii. To determine whether EC cases are increasing or decreasing for GPH over the out of sample period.

2.0 RELATED STUDIES

In a case-control study, Vizcaino *et al*. (1995) examined risk factors associated with EC in Bulawayo in Zimbabwe. Data was analyzed using a logistic regression model. The study basically found out that, in men, tobacco smoking was associated with EC. Based on statistical analysis and spatial econometrics, Boyko *et al*. (2015) investigated the epidemiology and forecasted the prevalence of EC in Central and Eastern Europe (Austria, Germany, Czech Republic, Poland Slovakia and Hungary). The study established that in all the countries under consideration, the prevalence of EC was high and was also projected to rise in the future. In an Indian study, Hasan *et al*. (2016) predicted the recurrence of free survival for EC patients using a protein signature based risk model and found out that biomarker signature score based on cytoplasmic beta-catenin, nuclear c-Myc, nuclear DVL and membrane alpha-catenin was associated with recurrence free survival. No study has attempted to model and forecast EC cases in the country. This paper is the first of its kind in Zimbabwe.

3.0 METHODOLOGY

In this paper we apply the Artificial Neural Network (ANN) approach in modeling and forecasting monthly EC cases at GPH. In line with Fischer & Gopal (1994), who argue that no strict rules exist for the determination of the ANN structure; the study applies the popular ANN (12, 12, 1) model based on the hyperbolic tangent activation function.

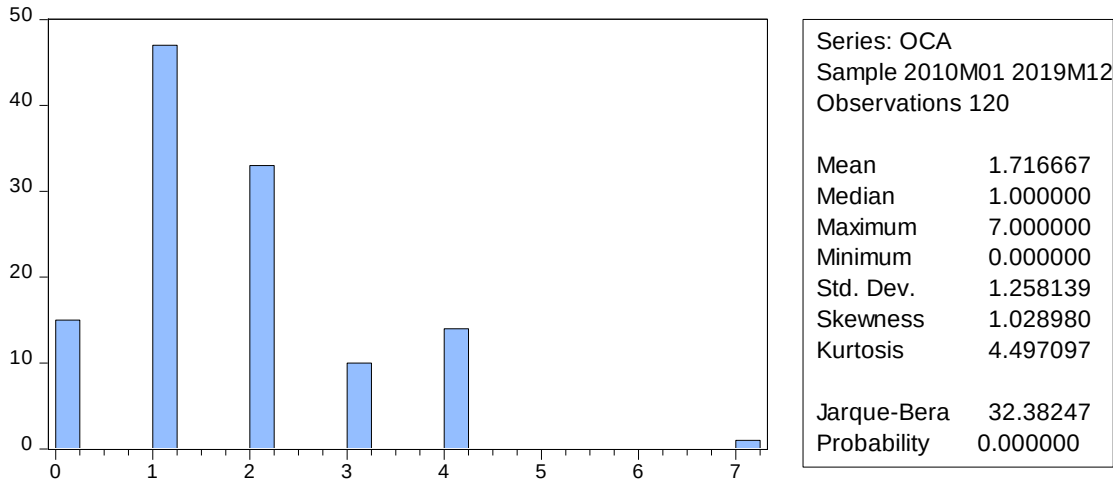
3.1 Data Issues

This study is based on newly diagnosed monthly EC cases (reffered to as OCA series in this study) in adults aged 45 years and above at GPH. 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 GPH Health Information Department.

4.0 FINDINGS OF THE STUDY

4.1 DESCRIPTIVE STATISTICS

Figure 1: Descriptive statistics



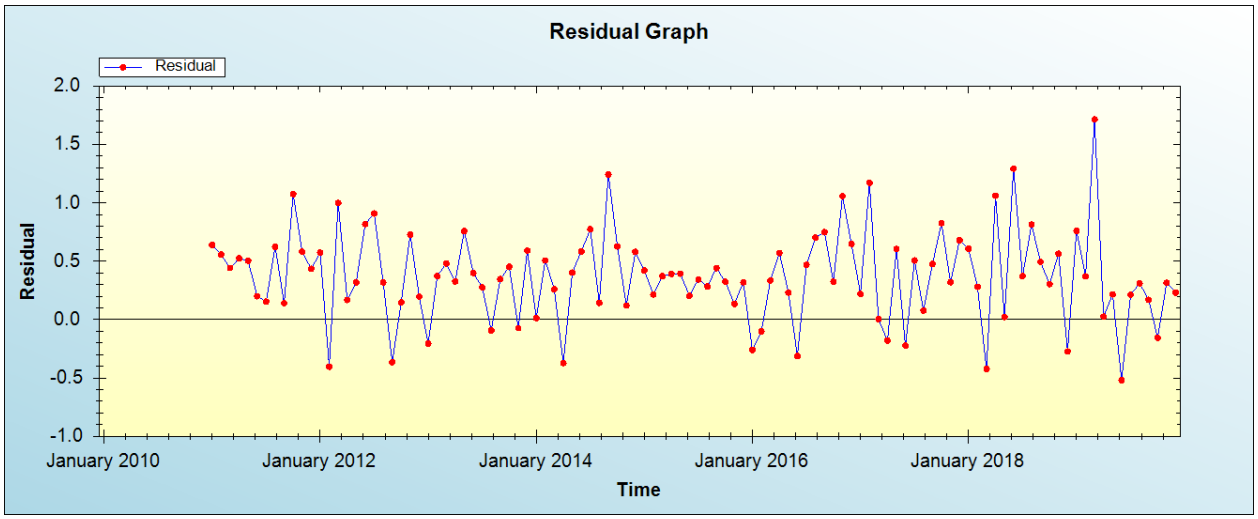
4.2 ANN Model Summary

Table 1: ANN model summary

Variable	OCA
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	0.138239
MSE	0.289010
MAE	0.446488

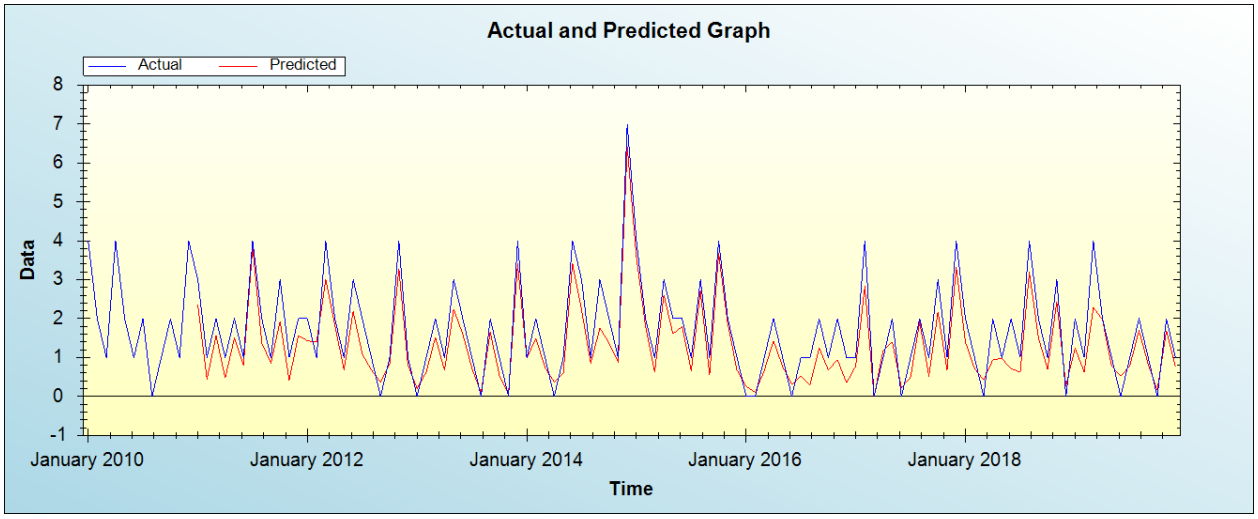
Residual Analysis for the Applied Model

Figure 2: Residual analysis



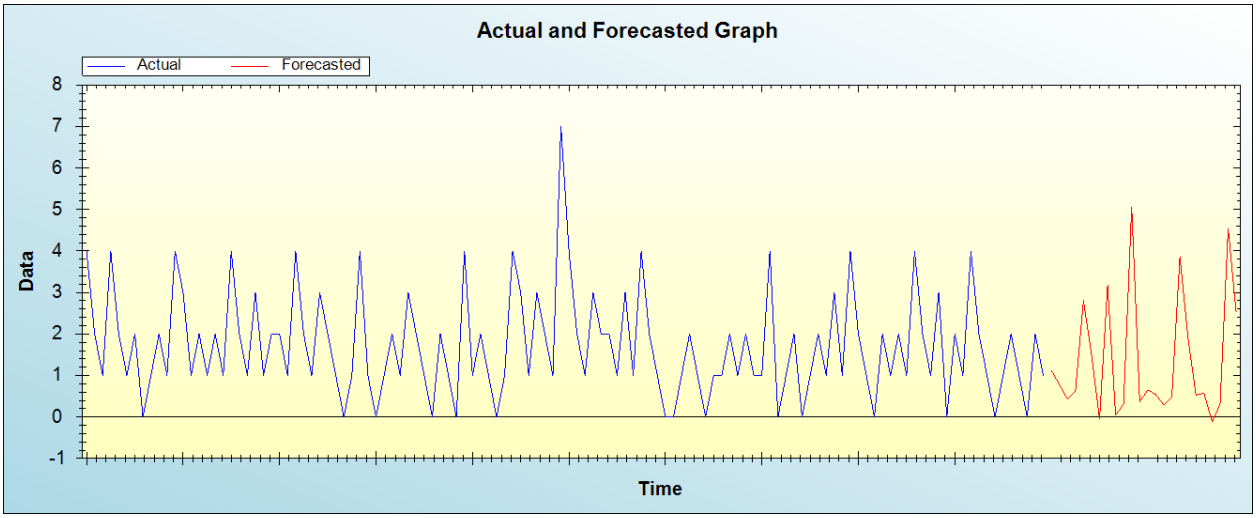
In-sample Forecast for OCA

Figure 3: In-sample forecast for the OCA series



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

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

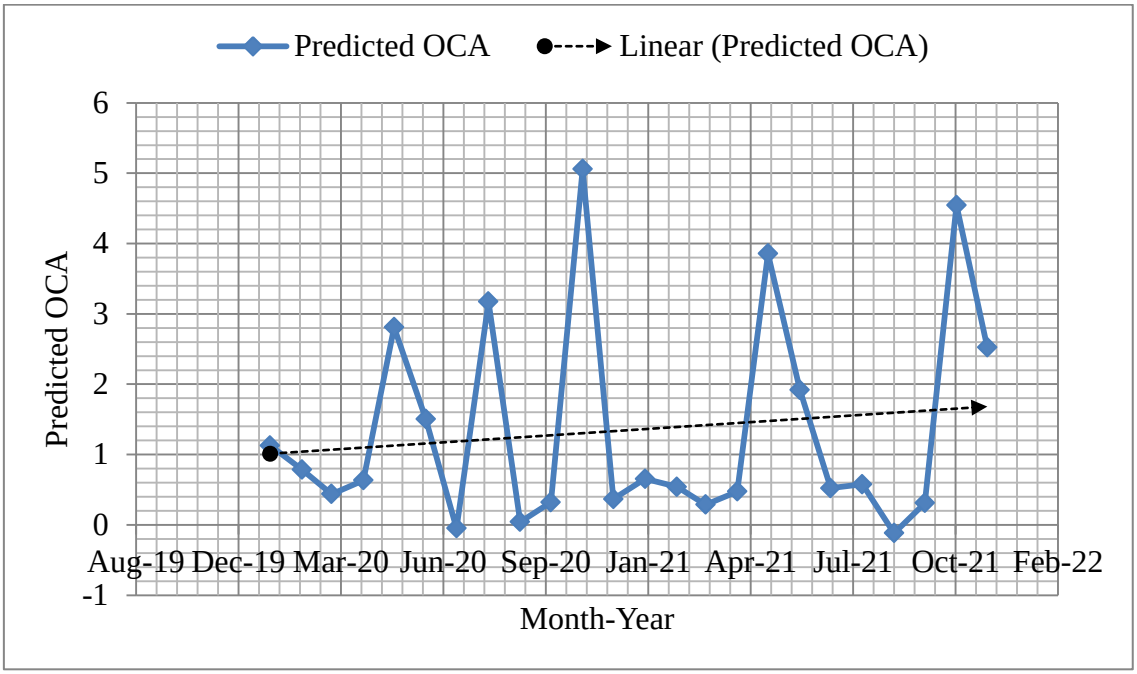


Out-of-Sample Forecast for OCA: Forecasts only

Table 2: Tabulated out-of-sample forecasts

Month/Year	Predicted OCA
January 2020	1.1269
February 2020	0.7862
March 2020	0.4379
April 2020	0.6323
May 2020	2.8079
June 2020	1.5059
July 2020	-0.0504
August 2020	3.1742
September 2020	0.0425
October 2020	0.3210
November 2020	5.0573
December 2020	0.3670
January 2021	0.6545
February 2021	0.5410
March 2021	0.2871
April 2021	0.4730
May 2021	3.8556
June 2021	1.9174
July 2021	0.5228
August 2021	0.5766
September 2021	-0.1187
October 2021	0.3095
November 2021	4.5447
December 2021	2.5247

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



4.3 DISCUSSION OF THE RESULTS

Table 1 is the summary of the ANN (12, 12, 1) model, which has been based on the hyperbolic tangent function as its activation function. The "criteria" are the evaluation statistics and they all point to the notion that the model is adequate. Figure 1, most importantly shows that, an average of 2 EC cases have been diagnosed at GPH over the period under study. Figure 2 shows the residuals of the applied model and since the residuals are as close to zero as possible, the model is reffered to as stable and acceptable for generating forecasts for GPH EC caseloads. Figure 3 shows the in-sample forecast of the model and it can be deduced that the model fits well with data. Figure 4, table 2 and figure 5 are out of sample predictions. The results of the study indicate that EC cases will generally rise over the out-of-sample period; characterised by possible seasonal peaks in the month of November each year.

5.0 CONCLUSION & POLICY DIRECTIONS

EC is an increasing concern due to poor prognosis, aggressive disease modalities, and a lack of efficient therapeutics (Tiasto *et al*., 2018). This study practically demonstrated the importance of applying an ANN model to estimate monthly EC caseloads as well as forecasting the future trend. Using time series data over the period Janaury 2010 to December 2019, the study accurately forecasted monthly EC case volumes over the out-of-sample period. The following policy directions are therefore suggested:

i. GPH needs to expand its EC screening services in order to facilitate early diagnosis and treatment.
ii. GPH should also consider conducting public awareness campaigns on EC and the need to prioritize proper diets and healthy life styles for communities in its catchment area. Since there is no effective treatment for this scourge, preventive measures such as these are highly recommended and could go a long in the fight against this "silent killer".
iii. The government of Zimbabwe should establish rehabilitative and palliative care facilities at GPH and other healthcare institutions in the country.

REFERENCES

1. Arnold, A., *et al*. (2012). Global Incidence of Oesophageal Cancer By Histological Subtype, *Gut*, 64: 387 – 394.
2. Boyko, V., *et al*. (2015). Epidemiology and Forecasting of the Prevalence of Esophageal Cancer in the Countries of Central and Eastern Europe, *Procedia Economics and Finance*, 24: 93 – 100.
3. Cancer Research (2017). Esophageal Cancer in the UK, *Cancer Research*, London.
4. Fischer, M. M., & Gopal, S. (1994). Artificial Neural Networks: A New Approach to Modeling Interregional Telecommunication Flows, *Journal of Regional Science*, 34 (4): 503 – 527.
5. GLOBOCAN (2018). Estimated Cancer Incidence, Mortality and Prevalence, *GLOBOCAN*, New York.
6. Hasan, R., *et al*. (2016). Prediction of Recurrence Free Survival for Esophageal Cancer Patients Using a Protein Signature Based Risk Model, *Oncotarget*, pp: 1 – 13.
7. Hou, X., *et al*. (2017). Non-coding RNAs: New Biomarkers and Therapeutic Targets for Esophageal Cancer, *Oncotarget*, 8: 43571 – 43578.

8. Kachala, R. (2010). Systematic Review: Epidemiology of Oesophageal Cancer in Sub-Saharan Africa, *Malawi Medical Journal*, 22: 65 – 70.
9. Middleton, D. R. S., *et al*. (2018). Esophageal Cancer – Male to Female Incidence Ratios in Africa: A Systematic Review and Meta – Analysis of Geographic, Time and Age Trends, *Cancer Epidemiology*, 53: 119 – 128.
10. Ministry of Health & Child Care (2014): National Cancer Prevention and Control Strategy for Zimbabwe: 2014 – 2018, *Government Printers*, Harare.
11. Palethorpe, H. M., Drew, P. A., & Smith, E. (2017). Androgen Signaling in Esophageal Adenocarcinoma Cell Lines, *Digestive Diseases Science*, 62: 3402 – 3414.
12. Rubenstein, J. H., & Shaheen, N. J. (2015). Epidemiology, Diagnosis and Management of Esophageal Adenocarcinoma, *Gastroenterology*, 149: 302 – 317.
13. Testa, U., Castelli, G., & Pelosi, E. (2017). Esophageal Cancer: Genomic Molecular Characterization, Stem Cell Compartment and Clonal Evolution, *Medicines*, 4: 1 – 12.
14. Tiasto, V., *et al*. (2018). Esophageal Cancer Research Today and Tomorrow: Lessons From Algae and Other Perspectives, *Genetics*, 5 (1): 75 – 90.
15. Vizcaino, A. P., Parkin, D. M., Skinner, M. E. G. (1995). Risk Factors Associated With Oesophageal Cancer in Bulawayo, Zimbabwe, *British Journal of Cancer*, 72: 769 – 773.
16. Wang, R. H. (2015). From Reflux Esophagitis to Barret's Esophagus and Esophageal Adenocarcinoma, *World Journal of Gastroenterology*, 21: 5210 – 5219.
17. Watanabe, M. (2015). Risk Factors and Molecular Mechanisms of Esophageal Cancer: Differences Between Histologic Subtypes, *Journal of Cancer Metastasis Treatment*, 1: 1 – 7.
18. Zhang, Y. (2013). Epidemiology of Esophageal Cancer, *World Journal of Gastroenterology*, 19: 5598 – 5606.