Multivariate Time Series Forecasting of Lung and Colon Cancer Mortality: A Comparative Study Between Jamaica and the United States

Shanice Douglas[1] Mostafa Zahed [2]

Department of Mathematics and Statistics, East Tennessee State University

Department of Mathematics and Statistics

Abstract

- This study utilizes **multivariate time series models** to investigate lung and colon cancer mortality trends in **Jamaica** and the **United States** (1960–2014). It explores interdependence between the two cancers and provides **15-years of forecasts** to support long-term cancer surveillance.
- Model accuracy is evaluated using **MAPE**, **RMSE**, and **AIC**. The results uncover regional disparities and shared trends, offering **data-driven insights** to strengthen cancer control strategies and public health policy.

Objectives and Motivation

- To forecast lung and colon cancer mortality in Jamaica and the United States using multivariate time series models.
- Lung and colon cancer are among the leading causes of cancer-related deaths globally.
- Jamaica's public health data has been underutilized in predictive modeling.
- Forecasting cancer mortality can support more informed health policy and resource allocation.
- Comparing cancer trends in Jamaica with those in the U.S. may reveal valuable insights for prevention and intervention.

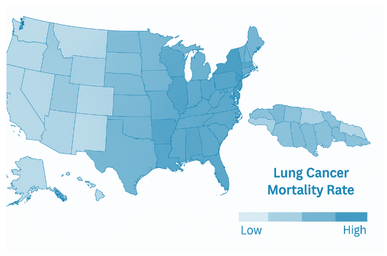


Figure 1. Geographic distribution of lung cancer mortality rates across the United States and Jamaica.

Introduction

- Lung and colon cancers are major contributors to global cancer mortality.
- Lung cancer is linked to smoking and environmental risks; colon cancer to diet and lifestyle.
- Both cancers are often detected late, especially in low-resource settings.
- The U.S. has screening programs, but mortality remains high; Jamaica faces growing cancer burdens.
- Time series modeling supports trend forecasting and public health decision-making.

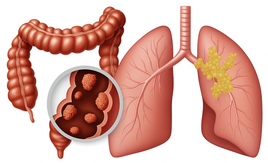


Figure 2. Visual illustration of lung and colon cancer.

United States vs Jamaica: Cancer Divide

Table 1. Comparison of Colon and Lung Cancer Outcomes and Access to Care in the United States and Jamaica

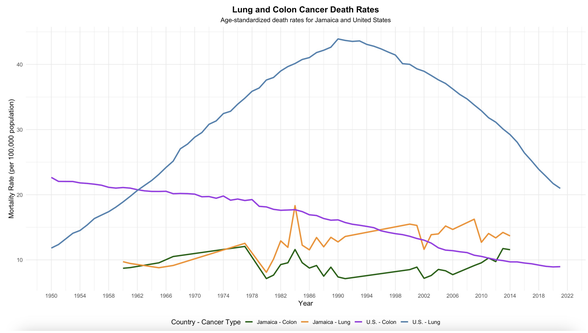
Indicator	United States	Jamaica
Stage at Diagnosis	Colon: 35% localized Lung: Often Stage 3 or 4	Colon: Mostly regional or distant Lung: Often diagnosed late
Survival Rate	Colon: 15.7% Lung: 6%	Colon: 33% overall Lung: Very low
Access to Care	Established screening and treatment services.	Limited screening, diagnostics, and treatment availability

Data Description

The dataset used for this analysis was obtained from Our World in Data, which provides globally curated cancer statistics. It includes age-standardized death rates per 100,000 individuals over several decades.

This project focuses on lung and colon cancer mortality trends in **Jamaica** and the **United States** from 1960 to 2014. Jamaica, a developing country with resource-limited healthcare infrastructure, is compared with the United States, a developed nation with more advanced medical and policy systems.

Below is the time series plot of lung and colon cancer death rates in Jamaica and the United States from 1950 to 2021.



Methodology

- Lung and colon cancer mortality data for Jamaica and the United States were compiled and organized into annual time series from 1960 to 2014.
- Stationarity of each series was assessed using the Augmented Dickey-Fuller (ADF) test, and differencing was applied where necessary.
- The Johansen cointegration test was used to evaluate long-term relationships between lung and colon cancer mortality within each country.
- A multivariate time series model, specifically a Vector Autoregressive Moving Average (VARMA) model, was fitted to capture interdependent trends.
- Model performance was evaluated using diagnostic metrics such as AIC, MAPE, and RMSE to ensure reliability and accuracy.
- Forecasts were generated and visualized to provide insight into the temporal dynamics of lung and colon cancer mortality across both countries.

Model Results

Fitted model for Lung and colon Cancer in United States VMA (1)

$$\begin{bmatrix} t_{1,t} \\ t_{2,t} \end{bmatrix} = \begin{bmatrix} -0.0156 \\ 0.0158 \end{bmatrix} + \begin{bmatrix} 0.6203 & -0.0051 \\ -0.0485 & 0.6973 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1,1} \\ \varepsilon_{t-1,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t,1} \\ \varepsilon_{t,2} \end{bmatrix} \quad (1)$$

Interpretation of VMA(1) Coefficients (United States)

Coefficient	Effect	Interpretation
Intercept (Lung)	-0.0156	Average change in lung cancer death rate is a slight decrease over time.
Intercept (Colon)	0.0158	Average change in colon cancer death rate is a slight increase over time.
MA(1): Lung → Lung	0.6203	Past shock in lung cancer rate strongly increases current lung rate.
MA(1): Colon → Lung	-0.0051	Past shock in colon cancer rate has a very small negative effect on lung rate.
MA(1): Lung → Colon	-0.0485	Past shock in lung cancer rate has a slight negative effect on colon rate.
MA(1): Colon → Colon	0.6973	Past shock in colon cancer rate strongly increases current colon rate.

Fitted model for Lung and colon Cancer in Jamaica VARMA(1,1)

$$\begin{bmatrix} z_{t1} \\ z_{t2} \end{bmatrix} = \begin{bmatrix} 0.060 \\ -0.004 \end{bmatrix} + \begin{bmatrix} 0.481 & -0.140 \\ 0.253 & 0.235 \end{bmatrix} \begin{bmatrix} z_{t-1,1} \\ z_{t-1,2} \end{bmatrix} + \begin{bmatrix} 0.980 & 0.393 \\ 0.591 & 0.341 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1,1} \\ \varepsilon_{t-1,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t,1} \\ \varepsilon_{t,2} \end{bmatrix} \quad (2)$$

Interpretation of VARMA(1,1) Coefficients (Jamaica)

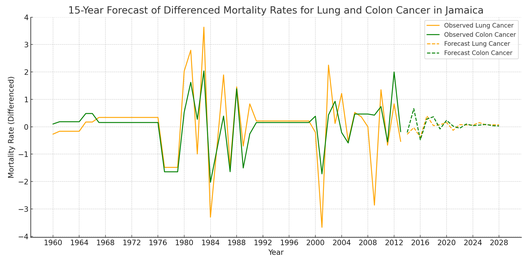
Coefficient	Effect	Interpretation
Intercept (Lung)	0.060	Lung cancer death rate slightly increases over time.
Intercept (Colon)	-0.004	Colon cancer death rate slightly decreases over time.
AR(1): Lung → Lung	0.481	Prior lung cancer changes moderately increase current lung rate.
AR(1): Colon → Lung	-0.140	Prior colon cancer changes slightly decrease lung rate.
AR(1): Lung → Colon	0.253	Prior lung cancer changes slightly increase colon rate.
AR(1): Colon → Colon	0.235	Prior colon cancer changes moderately increase colon rate.
MA(1): Lung → Lung	0.980	Past shocks in lung cancer strongly increase current lung rate.
MA(1): Colon → Lung	0.393	Past shocks in colon cancer moderately increase lung rate.
MA(1): Lung → Colon	0.591	Past shocks in lung cancer moderately increase colon rate.
MA(1): Colon → Colon	0.341	Past shocks in colon cancer modestly increase colon rate.

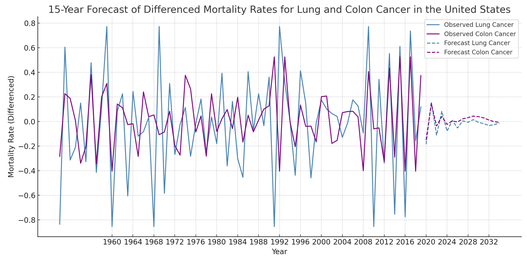
Forecast Accuracy Metrics

Table 2. Forecast Accuracy Metrics for Lung and Colon Cancer (Jamaica vs. USA)

Cancer Type & Country	RMSE	MAE	MAPE (%)
Lung Cancer (Jamaica)	1.078	0.945	6.68
Colon Cancer (Jamaica)	1.714	1.362	13.74
Lung Cancer (USA)	2.455	2.1934	6.16
Colon Cancer (USA)	4.8706	4.4706	23.18

Forecast results





Forecast models reveal contrasting trends in lung and colon cancer mortality across the United States and Jamaica, highlighting the importance of localized cancer surveillance.

Projected Change in Lung and Colon Cancer Mortality Rates (Year 1–15)

Table 3. Forecasted Annual Change in Lung and Colon Cancer Mortality Rates in the United States and Jamaica Over a 15-Year Period

Year	USA Lung	USA Colon	Jamaica Lung	Jamaica Colon
Year 1	-0.222	-0.036	-0.048	-0.028
Year 2	0.093	0.091	-0.081	0.017
Year 3	-0.118	-0.074	-0.070	0.014
Year 4	0.044	0.058	-0.062	0.010
Year 5	-0.061	-0.035	-0.080	0.007
Year 6	-0.004	0.043	-0.080	0.022
Year 7	-0.041	-0.024	-0.062	0.013
Year 8	-0.014	0.023	-0.070	0.015
Year 9	-0.027	-0.006	-0.082	0.007
Year 10	-0.024	0.014	-0.072	0.012
Year 11	-0.024	-0.001	-0.082	0.015
Year 12	-0.024	0.009	-0.082	0.009
Year 13	-0.023	0.004	-0.075	0.016
Year 14	-0.026	0.006	-0.097	0.011
Year 15	-0.023	0.005	-0.095	0.013

Conclusion and Future Work

- Multivariate time series models effectively captured cancer mortality trends and revealed disparities between Jamaica and the United States.
- Future work will compare modeling approaches to evaluate performance across different sample sizes and datasets.

Key References

- Kilian, L., & Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*. Cambridge University Press.
- Pascual, L., Romo, J., & Ruiz, E. (2004). Effects of parameter estimation on prediction densities: A case study with VAR models. *Computational Statistics & Data Analysis*, **47**(4), 675–698.

douglasss2@etsu.edu

zahedm@etsu.edu