Science World Journal Vol. 20(No 3) 2025 https://dx.doi.org/10.4314/swj.v20i3.15
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

MULTIVARIATE ANALYSIS OF VARIANCE APPLICATION TO CANCER-VARIANT MORTALITY RATES IN SELECTED WEST AFRICAN COUNTRIES

[1]Akarawak E.E.E.; *[1,2]Ikegwu Emmanuel M.

[1]Department of Statistics, University of Lagos, Akoka Lagos
[2]Department of Statistics, Yaba College of Technology, Yaba Lagos

*Corresponding Author Email Address: emmanuel.ikegwu@yabatech.edu.ng

ABSTRACT
The menace of cancer, which has sent many to untimely deaths globally in the 21st century, lives with us today. This study compared the reported death rates from five selected West African countries over nine different types of cancer, excluding all sexually related types. The data was acquired from Kaggle.com, covering the period from 1990 to 2019, and contained various cancer types; however, the study considered only nine different types from the five selected West African countries. The Multivariate Analysis of Variance (MANOVA) was adopted for the analysis and comparison of cancer-variant mortality rates. The results showed a significant difference in the mean vectors of the cancer-variant mortality rates, with Côte d'Ivoire consistently higher and Nigeria consistently lower ($p < 0.05$). It was therefore concluded that cancer-variant mortality rates in West African countries, as low-income countries, are still moderate and that efforts need to be made to reverse the rise observed in certain variants while improving on policies to keep those declining on the drop.

Keywords: Cancer-variants, MANOVA, Mortality rates, Pillai's test, West Africa

INTRODUCTION

Cancer has been reported as the leading cause of death globally and amongst the top three drivers of the increase in deaths from non-communicable diseases, which accounts for an estimated 37% of all deaths in Sub-Saharan Africa (SSA) (Mutebi, 2023; WHO, 2022). Cancer, according to the National Cancer Institute (2021), is an ailment in which "some of the body's cells grow uncontrollably and spread to other parts of the body". The World Health Organisation identified lung, prostate, colorectal, stomach, and liver cancer as the most common in men, and breast, colorectal, lung, cervical and thyroid as most common among women (WHO, 2022). Researchers have noted that "cancer fatalism" makes many cancer patients believe that the diagnosis is a sure death sentence and that treatment is futile, which has led to many deaths from different cancer infections even when care and treatment have improved (Mutebi, 2023; Duru & Topatan, 2023; Salisu *et al*., 2022; Powe & Finnie, 2003). Fatalism was defined by Keller *et al.* (2021) as an "emotional state of gloom, hopelessness, and helplessness concerning cancer outcomes. Fatalism is driven by cultural and social beliefs and has raised the rates of death from cancer in Africa, coupled with low health access, out-of-pocket treatment and general poverty (Alberto *et al.,* 2023; Kim & Lwin, 2021).

The implementation of multivariate models in the study of health-related phenomena has persisted because of the complex nature of human health and its interrelatedness within the environment. Kandula *et al.* (2023) used a hybrid classification model (a voting classifier) to study personalised cancer with gene, variant, and text features. Hancock (2006) employed the multivariate consensus tree, a tree-based clustering for mixed data types on thyroids. Mo *et al*. (2021) used a comprehensive transcriptomic analysis on an EMT-related gene signature in colorectal cancer. Yanai *et al.* (1979) employed factor analysis to analyse cancer mortalities in 24 selected countries from Africa, Europe, America, Australia, and Asia. Rubio *et al*. (2023) used exploratory analysis and modelling to study the geospatial distribution of breast cancer mortality rate in Colombia, and also used logistic regression, K-nearest neighbour, support vector machines, naïve Bayes, decision trees, and random and rotation forest, multivariate methods, to analyse and compare Wisconsin breast cancer detection and diagnostics.

This study aims to compare the proportion of deaths from cancer variants in selected West African countries that share a common dip into the Atlantic Ocean using the Multivariate Analysis of Variance (MANOVA). The specific objectives included obtaining the mean vector and variance-covariance matrix of cancer-variants mortality rates (CVMR) for the countries under study, and comparing the mean vectors using the multivariate analysis of variance of the cancer-variants mortality rates.

METHODOLOGY

The data used for the study were mortality proportions of 30 years (1990 – 2019) for 9 different cancer types (Bladder, Kidney, Stomach, Esophageal, Tracheal, bronchus and lung (TBL), Colon and rectum, Gallbladder and biliary tract, Liver, and Brain and central nervous system) collected from 5 west African countries (Cote d'Ivoire, Ghana, Liberia, Nigeria and Senegal) who had the foot in the Atlantic Ocean.

Model Specifications

The Multivariate analysis of variance (MANOVA) is an extension of the univariate Analysis of variance for inferences when the population groups are more than two ($k > 2$). The Multivariate Analysis of variance has as assumptions that the observations are independent of each other, the dependent variables are multivariate normally distributed, and the population covariance matrix is equal, that is, homogeneity of variance.

The multivariate linear model with p >1 response variables is an extension of the univariate model with added columns to cater for the dependent variables and additional regression coefficients associated with each of them, and additional columns for the random error associated with each additional response variable. The Multivariate model, according to Haase (2012), is specified as:

Science World Journal Vol. 20(No 3) 2025 https://dx.doi.org/10.4314/swj.v20i3.15
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

$$\boldsymbol{Y}_{(n\times p)} = \boldsymbol{X}_{(n\times(q+1))}\boldsymbol{B}_{((n-1)\times p)} + \boldsymbol{E}_{(n\times p)} \quad (1)$$

Where **Y is an** $n \times p$ matrix of response variables, **B** is an $(n-1) \times p$ matrix of regression coefficients, **X** is an $n \times (q-1)$ matrix of predictor variables, and **E** is an $n \times p$ matrix of the disturbances or random errors. It is noteworthy that while the design matrix **X** is comparable to the univariate case, the multiplicity of the response variables, estimates of the parameters, and the disturbances that characterise the multivariate linear model need to be developed. Jaccard & Jacoby (2010) identified the choice of reliable and valid criterion and predictor variables hinged on the theoretical description of the hypothesised relationships (concepts, magnitude and direction) and specified models consistent with theories as building blocks for modelling.

Specifying the multivariate linear model commences with defining the four matrices of Eqn. 1 above with p > 1. The dependent variable $\boldsymbol{Y}_{(n\times p)}$ is an n x p matrix given as;

$$\boldsymbol{Y}_{(n\times p)} = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{bmatrix} \quad (2)$$

Also, the explanatory variables of the model, which make up the design matrix $\boldsymbol{X}_{(n\times(q+1))}$, consist of the q predictor measures $(X_1, X_2, \ldots, X_q)$ and the unit column vector $X_0 = 1$ used for estimating the model intercept. The design matrix has a general form of:

$$\boldsymbol{X}_{(n\times(n\pm1))} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \quad (3)$$

In the matrix of model parameters **B** in equation 1, the multiple dependent variables are accompanied by multiple columns of **B** to accommodate all of the **Y-X** relationships with an order of q + 1. The rows of B correspond to the predictor variables $(X_0, X_1, X_2, \ldots, X_q)$ and the columns represent the response variables $(Y_1, Y_2, \ldots, Y_p)$ and the model is given as:

$$\boldsymbol{B}_{((q\pm1)\times p)} = \begin{bmatrix} \beta_{01} & Y_{02} & \cdots & \beta_{0p} \\ \beta_{11} & Y_{12} & \cdots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{bmatrix} \quad (4)$$

The matrix product $\boldsymbol{X}_{(n\times(n\pm1))}\boldsymbol{B}_{((q+1)\times p)}$ in equation 1 conforms multiplicatively while the order of $\boldsymbol{XB}_{(n\times p)}$ is determined by **X** rows and **B** columns and given as:

$$\boldsymbol{XB}_{(n\times p)} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_{01} & Y_{02} & \cdots & \beta_{0p} \\ \beta_{11} & Y_{12} & \cdots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{bmatrix} \quad (5)$$

Using the additive equality, equations (2) to (5) yield:

$$\begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_{01} & Y_{02} & \cdots & \beta_{0p} \\ \beta_{11} & Y_{12} & \cdots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{bmatrix} \pm \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{np} \end{bmatrix} \quad (6)$$

Which is equivalent to equation 1 and reveals that $\boldsymbol{XB}_{(n\times p)}$ conforms to the order of $\boldsymbol{E}_{(n\times p)}$ which satisfies the order of $\boldsymbol{Y}_{(n\times p)}$.

MANOVA Computations

The multivariate model null hypothesis of equality of means-tested, according to Johnson & Wichern (2014), is given as

$$H_0: \tau_1 = \tau_2 = \cdots = \tau_p = \sum_{l=1}^{p} n_l\tau_l = 0 \quad (7)$$

The sums of squares and cross-products are expressed as:

$$\sum_{l=1}^{p}\sum_{j=1}^{n_l}(x_{lj} - \bar{x})(x_{lj} - \bar{x})' = \sum_{l=1}^{p} n_l(\bar{x}_l - \bar{x})(\bar{x}_l - \bar{x})' + \sum_{l=1}^{p}\sum_{j=1}^{n_l}(x_{lj} - \bar{x}_l)(x_{lj} - \bar{x}_l)' \quad (8)$$

$$\begin{pmatrix} total\ (corrected)\ sum \\ of\ squares\ and\ cross \\ products \end{pmatrix} = \begin{pmatrix} treatment\ (\boldsymbol{B}etween)\ sum \\ of\ squares\ and\ cross \\ products \end{pmatrix} + \begin{pmatrix} residual\ (\boldsymbol{W}ithin)\ sum \\ of\ squares\ and\ cross \\ products \end{pmatrix}$$

Where the within sum of squares and cross products matrix is written as

$$\boldsymbol{W} = \sum_{l=1}^{p}\sum_{j=1}^{n_l}(x_{lj} - \bar{x}_l)(x_{lj} - \bar{x}_l)' = (n_1 - 1)\boldsymbol{S}_1 \pm (n_2 - 1)\boldsymbol{S}_2 \pm \cdots \pm (n_p - 1)\boldsymbol{S}_p \quad (9)$$

Where $\boldsymbol{S}_l$ is the sample covariance matrix for the lth sample.

The between sums of squares and cross product is equally expressed as

$$\boldsymbol{B} = \sum_{l=1}^{p} n_l(\bar{x}_l - \bar{x})(\bar{x}_l - \bar{x})' \quad (10)$$

The multivariate computations that lead to the test statistic are summarised in the MANOVA Table 1.

Science World Journal Vol. 20(No 3) 2025 https://dx.doi.org/10.4314/swj.v20i3.15
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Table 1: MANOVA table for comparing Population Mean Vectors

Source of variation	Matrix of the sum of squares and cross products (SSP)	Degrees of freedom (df)
Treatment	$\boldsymbol{B} = \sum_{l=1}^{p} n_l(\bar{x}_l - \bar{x})(\bar{x}_l - \bar{x})'$	$p-1$
Residual (Error)	$\boldsymbol{W} = \sum_{l=1}^{p}\sum_{j=1}^{n_l}(x_{lj} - \bar{x}_l)(x_{lj} - \bar{x}_l)'$	$\sum_{l=1}^{p} n_l - p$
Total (corrected for the mean)	$\boldsymbol{B} \pm \boldsymbol{W} = \sum_{l=1}^{p}\sum_{j=1}^{n_l}(x_{lj} - \bar{x})(x_{lj} - \bar{x})'$	$\sum_{l=1}^{p} n_l - 1$

Source: Johnson & Wichern (2014)

Multivariate Statistical tests

The MANOVA model with k levels of independent variables and p number of dependent variables has several tests which are stated in terms of between-group (**B**) and within-group (**W**) variances which depict matrices of sums of squares of Y and their cross-products (that is, the variances and covariances matrices that are not divided by the b and w (where w > b) degrees of freedom (df). These are tested against the F-distributions for significance. Significant results imply differences among the groups on dependent variables taken together.

Wilks' Lambda

The Wilks' Lambda measures the ratio of the mean square within to the mean square total.

$$\boldsymbol{\Lambda} = \frac{|\boldsymbol{W}|}{|\boldsymbol{B} \pm \boldsymbol{W}|} = \frac{|\boldsymbol{W}|}{|\boldsymbol{T}|} \quad (11)$$

Mertler & Reinhart (2017) observed that Wilks' $\boldsymbol{\Lambda}$ is an "inverse criterion", implying that the treatment effects or group is significantly different when the Lambda statistic is smaller.

The effect size of the treatments or groups is computed with the eta squared ($\boldsymbol{\eta^2}$), which is the variance-covariance accounted for by the best combination of the multiple dependent variables (Hahs-Vaughn, 2017; Tabachnick & Fidell, 2019) and is computed as:

$$\boldsymbol{\eta^2} = \mathbf{1} - \boldsymbol{\Lambda} \quad (12)$$

The F-distribution approximation with the determined significance levels is given as

$$\boldsymbol{F_{pb,ft-w}} = \frac{(ft-c)(1-\Lambda^{1/t})}{\mathrm{ph}\Lambda^{1/t}} \quad (13)$$

where

$$f = w - \frac{1}{2}(p - b \pm 1)$$

$$c = \frac{1}{2}pb - 2$$

$$t = \begin{cases} \sqrt{\frac{p^2b^2-4}{p^2+b^2-5}}, & if\ p^2 + b^2 - 5 > 0 \\ 1 & otherwise \end{cases}$$

The approximation is exact if and only if p or b ≥ 2.

Lawley - Hotelling's Trace

Hotelling's Trace or Lawley-Hotelling trace is a generalisation of Hotelling's T^2, applied to k > 2 groups. Trace originates from the matrix function that sums the diagonals of the matrix, that is, the sum of the variances in the variance-covariance matrix. It is given as:

$$\frac{T^2}{N-k} or\ trace(\boldsymbol{W}^{-1}\boldsymbol{B}) \quad (14)$$

where N is the total sample size, k is the number of groups, **B** is the matrix of the sum of squares cross-products for the hypothesis (explained), and **W** is the matrix sum of squares cross-products of errors.

The F-distribution approximation with the determined significance levels is given as

$$F_{a,i} = \frac{T_c^2}{gw} \quad (15)$$

where

$$a = pb \quad (16)$$

$$i = 4 + \frac{a+2}{D-1} \quad (17)$$

$$g = \frac{a(i-2)}{i(w-p-1)} \quad (18)$$

$$D = \frac{(w+b-p-1)(w-1)}{(w-p-3)(w-p)} \quad (19)$$

Where D defines the dimensionality or number of variables in the multivariate problem, a and b are the parameters, which relate to the number of degrees of freedom of complex Wishart distributions, g is the degree of freedom for the number of looks derived from the data, and i is the index of individual observations.

Pillai's Trace

The Pillai's trace is given as:

$$V^{(s)} = trace[\boldsymbol{B}(\boldsymbol{B}+\boldsymbol{W})^{-1}] \quad (20)$$

The F-distribution approximation with the determined significance levels is given as

$$F_{s(2m+s+1),2n\pm s\pm 1)} = \frac{(2n+s+1)V^{(s)}}{(2m+s+1)(s-V^{(s)})} \quad (21)$$

where

$$s = \min(p, b) \quad (22)$$

$$m = 0.5\,(|p-b| - 1) \quad (23)$$

$$n = 0.5\,(w-p-1) \quad (24)$$

Roy's Largest Root

Roy's largest root or Roy's greatest root, φ_{max}, is the largest eigenvalue of the matrix $\boldsymbol{W}^{-1}\boldsymbol{B}$

$$\varphi_{max} = \max(\varphi_i) \quad (25)$$

$$\varphi_i = \boldsymbol{W}^{-1}\boldsymbol{B} \quad (26)$$

The F-distribution approximation with the determined significance levels is given as

$$F_{(2v_1\pm 2),(2v_2\pm 2)} = \frac{2v_2 \pm 2}{2v_1 \pm 2}\varphi_{max} \quad (27)$$

where

$$s = \min(p, b) \quad (28)$$

$v_1 = m$ in equation (23)

$v_2 = n$ in equation (24)

Science World Journal Vol. 20(No 3) 2025 https://dx.doi.org/10.4314/swj.v20i3.15
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

The Pillai's Trace, Wilks' Lambda, Hotelling's Trace, and Roy's largest root are all equal when k = 2. While Roy's largest root is too probable to yield Type I errors, Wilks' lambda and Hotelling's trace are sensitive to the violations of homogeneity of covariances with small samples, and Pillai's trace is recommended for general use.
Box's M test of equality of covariances is used in multivariate analysis to test whether the covariance matrices of the dependent variables are equal across the groups formed by the independent variables.
The data collected from 1990 – 2019 for Bladder, Kidney, Stomach, Oesophagal, Tracheal, bronchus, and lung (TBL), Colon and rectum, Gallbladder and biliary tract (GBT), Liver and Brain and central nervous system (BCNS) cancers sourced from data.worldbank.org/database. The data collected were analysed with Multivariate Analysis of Variance, Univariate Analysis of Variance and the post hoc test was by the Duncan Multiple Range Test (DMRT) (Midway *et a.*, 2020), with the aid of Statistical Software for Social Sciences (SPSS 27.0) and R.

RESULTS

The results include tables and Multivariate statistics computed from the cancer variants' mortality rates.

Table 2: Box's M test of equality of Covariances for CVMR

Box's Test of Equality of Covariance Matrices	
Box's M	3252.149
F	15.639
df1	180
df2	37801.798
Sig.	.000
Tests the null hypothesis that the observed covariance matrices of z, the dependent variables, are equal across groups.	

The Box's M test in Table 2 gave a **χ^2 of 3252.149,** which is significant ($p < 0.05$), implying that the null hypothesis is rejected and the assumption of equal covariance across the groups (cancer variants) is violated. Hence, the use of Wilks' statistic is misleading, leaving us with the more robust Pillai's Trace test statistic, which is more robust to the violation of the assumption.

Table 3: Multivariate Tests of the CVMR

Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	Pillai's Trace	1.000	25859.558	9.000	108.000	<.000	1.000
	Wilks' Lambda	4.64E-4	25859.558	9.000	108.000	<.000	1.000
	Hotelling's Trace	2154.963	25859.558	9.000	108.000	<.000	1.000
	Roy's Largest Root	2154.963	25859.558	9.000	108.000	<.000	1.000
Country	Pillai's Trace	3.890	437.348	36.000	444.000	.000	.973
	Wilks' Lambda	3.799E-9	1982.404	36.000	406.463	.000	.992
	Hotelling's Trace	1784.279	5278.491	36.000	426.000	.000	.998
	Roy's Largest Root	1266.306	15617.779	9.000	111.000	<.000	.999
Year	Pillai's Trace	2.675	1.692	261.000	1044.000	<.000	.297
	Wilks' Lambda	.001	4.236	261.000	952.150	<.000	.522
	Hotelling's Trace	48.263	19.642	261.000	956.000	<.000	.843
	Roy's Largest Root	41.943	167.772	29.000	116.000	<.000	.977

Table 3 shows the multivariate tests where each statistic tests the hypothesis that the mean vectors of the cancer-variants mortality rates (CVMR) differ across the countries and over the years. The Pillai's Trace with a value of 3.890 means that almost 96.11% of the variance in the combination of all the cancer variants' mortality rates is explained by the country differences, and these differences are significant ($p < 0.05$), and this is in agreement with other statistics for the selected countries. Also, Pillai's Trace with a value of 2.675 means that almost 97.425% of the variance in the combination of all the cancer variants' mortality rates is explained by the year differences, and these differences are significant ($p < 0.05$), and this is in agreement with other statistics for the selected countries.
The Partial Eta Squared (η^2) of 0.973 shows that 97.3% of the variance in all the cancer variant mortality rates was relatively explained by the countries, while 29.7% was relatively explained by the year.

Science World Journal Vol. 20(No 3) 2025 https://dx.doi.org/10.4314/swj.v20i3.15
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Table 4: Mean ± Standard deviation of the cancer mortality rates (per 100,000) of some West African Countries

Cancer-Variants	CIV	GHA	LIB	NGA	SEN	F-Statistics	K-S (p)
Bladder cancer	3.79 ± 0.61[a]	3.16 ± 0.36[d]	3.57 ± 0.63[b]	1.16 ± 0.08[e]	3.39 ± 0.42[c]	2533.986 (<0.001)	0.180 (<0.001)
Kidney cancer	0.98 ± 0.06[a]	0.95 ± 0.06[b]	0.77 ± 0.11[d]	0.74 ± 0.08[e]	0.83 ± 0.07[c]	693.036 (<0.001)	0.084 (<0.001)
Stomach cancer	17.28 ± 1.27[a]	10.16 ± 1.23[d]	13.99 ± 1.21[c]	4.93 ± 0.18[e]	14.50 ± 1.05[b]	2605.425 (<0.001)	0.145 (<0.001)
Esophageal cancer	4.94 ± 0.47[a]	3.46 ± 0.17[d]	4.39 ± 0.67[b]	1.03 ± 0.05[e]	4.20 ± 0.59[c]	1182.90 (<0.001)	0.190 (<0.001)
TBL cancer	13.55 ± 0.36[a]	8.91 ± 0.27[d]	9.49 ± 0.69[c]	7.74 ± 0.71[e]	11.84 ± 0.99[b]	851.99 (<0.001)	0.152 (<0.001)
Colon and rectum cancer	9.36 ± 0.22[a]	7.51 ± 0.78[c]	6.16 ± 0.55[d]	7.34 ± 0.99[c]	7.82 ± 0.50[b]	475.73 (<0.001)	0.097 (<0.001)
Gallbladder and biliary tract cancer	1.40 ± 0.15[a]	1.29 ± 0.10[b]	1.24 ± 0.15[c]	0.82 ± 0.03[d]	1.26 ± 0.11[bc]	2331.73 (<0.001)	0.173 (<0.001)
Liver cancer	7.07 ± 1.59[a]	6.43 ± 0.35[b]	6.30 ± 0.76[b]	3.62 ± 0.06[c]	2.26 ± 0.08[d]	681.422 (<0.001)	0.151 (<0.001)
Brain and CNS cancer	1.19 ± 0.19[c]	2.44 ± 0.07[a]	1.04 ± 0.15[d]	1.54 ± 0.12[b]	1.07 ± 0.22[d]	2515.325 (<0.001)	0.177 (<0.001)

Countries with the same superscripts across each cancer variant are not significantly different at 5%.

Having shown that the cancer variants' mortality rates differ significantly across the countries in Table 3, univariate Analysis of Variance was conducted for the cancer variants. Table 4 shows the mean and standard deviation of the cancer mortality rates for the different types of cancer recorded in the selected West African Countries and the univariate Analysis of Variance for each cancer variant. The Table 4 show that Cote d'Ivoire has the highest mortality rate of 3.79 (sd = 0.61) for Bladder cancer, while Nigeria has the lowest mortality rate of 1.16 (sd = 0.08), and the Bladder cancer mortality rates across the five West African countries are significantly different from each other ($p < 0.05$). Also, it shows that Cote d'Ivoire has the highest mortality rate of 0.98 (sd = 0.06) for Kidney cancer, while Nigeria has the lowest mortality rate of 0.74 (sd = 0.08), and the Kidney cancer mortality rates across the five West African countries are significantly different from each other ($p < 0.05$).

In addition, it revealed that Cote d'Ivoire has the highest mortality rate of 17.28 (sd = 1.27) for Stomach cancer, while Nigeria has the lowest mortality rate of 4.93 (sd = 0.18), and the Stomach cancer mortality rates across the five West African countries are significantly different from each other ($p < 0.05$). Also, it shows that Cote d'Ivoire has the highest mortality rate of 4.94 (sd = 0.47) for oesophageal cancer, while Nigeria has the lowest mortality rate of 1.03 (sd = 0.05), and the oesophageal cancer mortality rates across the five West African countries are significantly different from each other ($p < 0.05$).

Furthermore, the Table 4 revealed that Cote d'Ivoire has the highest mortality rate of 13.55 (sd = 0.36) for TBL cancer, while Nigeria has the lowest mortality rate of 7.74 (sd = 0.71), and the TBL cancer mortality rates across the five West African countries are significantly different from each other ($p < 0.05$). Also, it shows that Cote d'Ivoire has the highest mortality rate of 9.36 (sd = 0.22) for Colon and rectum cancer, while Liberia has the lowest mortality rates of 6.16 (sd = 0.55) and the Colon and rectum cancer mortality rates across the five West African countries is significantly different from each other ($p < 0.05$).

Likewise, it revealed that Cote d'Ivoire has the highest mortality rate of 1.40 (sd = 0.15) for Gallbladder and biliary tract cancer while Nigeria has the lowest mortality rates of 0.82 (sd = 0.03) and the Gallbladder and biliary tract cancer mortality rates across the five West African countries significantly differs from each other ($p < 0.05$). Also, it shows that Cote d'Ivoire has the highest mortality rate of 7.07 (sd = 1.59) for Liver cancer, while Senegal has the lowest mortality rate of 2.26 (sd = 0.08), and the Liver cancer mortality rates across the five West African countries are significantly different from each other ($p < 0.05$).

Lastly, it revealed that Ghana has the highest mortality rate of 2.44 (sd = 0.07) for Brain and CNS cancer, while Liberia has the lowest mortality rate of 1.04 (sd = 0.15), and the Brain and CNS cancer mortality rates across the five West African countries was significantly different from each other ($p < 0.05$).

In general, it shows that Cote d'Ivoire shows consistently higher cancer mortality rates, while Nigeria shows lower cancer mortality rates, and the other three countries remain in between them. However, the Kolmogorov-Smirnov test shows that the data for none of the cancer variants is normally distributed even with transformation.

DISCUSSION

The study found various mortality rates range for the nine cancer variants across the different countries, 1.16 – 3.79 for Bladder cancer, 0.74 – 0.89 for Kidney cancer, Stomach cancer (4.93 – 17.28), Esophageal (1.03 – 4.94), Tracheal, bronchus and lung cancer (7.74 – 13.55), Colon and rectum cancer (6.16 – 9.36), Gallbladder and biliary tract cancer (0.82 – 1.40), Liver cancer (2.26 – 7.07) and Brain and central nervous system (1.04 – 2.44). These results were in tandem with the mean mortality rates reported by (Bray *et al*., 2024; Sharma *et al.,* 2022; Hamdi *et al*., 2021), while others were different from the global means. The mortality rate of bladder cancer in West African countries was found to be lower than the African average of over 8 per 100,000 and about 3.5 per 100,000 in northern Africa (Pizzato *et al.,* 2024; Morgan *et al*., 2022; Adeloye *et al*., 2019).

The study also found that the assumption of normality and equal variance-covariance across the different cancer variants was violated, but using the Robust Pillai's Trace statistic, the multivariate test was interpreted to reach conclusions. However, care must be taken in the conclusions on the results that follow, as

Liu, *et al.,* (2023) observed that none of the multivariate tests was robust. Though Olson (1976) observed that it is more important to watch out for type I and type II errors in the choice of test adopted, there has not been agreement or preference on any of the tests as in most cases these assumptions are violated even with simulated data under various scenarios and it was not different in the real-life data used for this study as transformation did not help handle these violations (Ates*, et al.,* (2019); Adeleke*, et al*., (2015)).

The study further found significant multivariate results in the mortality rates across the different cancer variants for the selected West African countries and across the 30 years studied. The subsequent univariate Analysis of Variance test shows that Cote d'Ivoire has a consistently higher mortality rate for the cancer variants studied, while Nigeria consistently had the lowest mortality rates from the cancer variants. The consistency in ranking may not have any real import, as it has been observed that incomplete recording, non-reporting of cases, inadequate infrastructure for capturing cancer events and weak policies and political will in dealing with cancer-related issues may actually have contributed to the variations found. There was no study found to compare our results because there have not been studies specific on the mortality rate of various variants of cancer but most studies have measured crude rate for all cancer variants but not variant-specific and others have looked at the percentage for all cancer deaths by countries (Igbokwe, (2023); Smailyte & Kurtinaitis, (2008); Kamangar, *et al.,* (2006)).

Conclusion

Having compared the CVMR of five West African countries from 1990 to 2019 using the multivariate analysis of variance (MANOVA), the study concluded that over the study periods, there was a significant difference in the vector of means of the cancer-variants for the five studied countries, with Nigeria consistently lower and Cote d'Ivoire consistently higher rates. Care should be taken in digesting these results without question, as many challenges have bedeviled data capturing and reporting, especially in these low-income countries (LICs). However, with violations of the normality and homogeneity of variance-covariance matrix assumptions of MANOVA by the real-life data, the use of the results should be with caution and efforts put in place to evolve methods that will correct for the violations.

Conflict of interest: The authors declare no conflict of interest in the study

Funding: The authors declare that the study was funded out of pocket

REFERENCES

Adeleke, B. I., Yahya, W. B., & Usman, A. (2015). A comparison of some test statistics for multivariate analysis of variance model with non-normal responses. *Journal of Natural Sciences Research, 5*(15), 1 – 9.

Adeloye, D., Harhay, M. O., Ayepola, O. O., Dos Santos, J. P., David, R. A., Ogunlana, O. O., . . . Rebbeck, T. R. (2019). An estimate of the incidence of bladder cancer in Africa: a systematic review and Bayesian meta-analysis. *International Journal of Urology, 26*(1), 102 - 112. doi:10.1111/iju.13824.

Alberto, N. R., Alberto, I. R., Puyat, C. V., Antonio, M. A., Ho, F. D., Dee, E. C., . . . Eala, M. A. (2023). Disparities in access to cancer diagnostics in ASEAN member countries. *The Lancet Regional Health – Western Pacific, 6*(32), 100667. doi:10.1016/j.lanwpc.2022.100667

Ates, C., Kaymaz, O., Kale, H. E., & Tekindal, M. A. (2019). Comparison of Test Statistics of Nonnormal and Unbalanced Samples for Multivariate Analysis of Variance in terms of Type-I Error Rates. *Computational and Mathematical Methods in Medicine, 2019*, 1 - 8. doi:10.1155/2019/2173638

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjamataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 74*(3), 229 - 263. doi:10.3322/caac.21834

Duru, G., & Topatan, S. (2023). A barrier to participation in cervical cancer screenings: fatalism. *Women's Health, 63*(6), 436 - 444. doi:10.1080/03630242.2023.2223698

Haase, R F. (2012). *Multivariate General Linear Models.* New York: SAGE Publications, Inc.

Hahs-Vaughn, D. L. (2017). *Applied Multivariate Statistical Concepts.* New York: Taylor & Francis.

Hamdi, Y., Abdeljaoued-Tej, I., Zatchi, A. A., Abdelhak, S., Boubaker, S., Brown, J. S., & Benkahla, A. (2021). Cancer in Africa: The Untold Story. *Frontiers of Oncology, 11*, 650117. doi:10.3389/fonc.2021.650117

Hancock, T. P. (2006). *Multivariate Consensus Tree: Tree-Based Clustering and Profiling for Mixed Data Types.* Australia: James Cook University. Retrieved June 29, 2024, from http://eprints.jcu.edu.au/17497

Igbokwe, K. K. (2023). Comparative examination of breast cancer burden in sub-Saharan Africa, 1990–2019: estimates from Global Burden of Disease 2019 study. *Oncology, 14*(3), 082492. doi:10.1136/bmjopen-2023-082492

Jaccard, J., & Jacoby, J. (2010). *Theory Construction and Model-Building Skills: A Practical Guide for Social Scientists.* New York: Guilford Press.

Johnson, R., & Wichern, D. (2014). *Applied Multivariate Statistical Analysis* (6th ed.). Harlow: Pearson Education Limited.

Kamangar, F., Dores, G. M., & Anderson, W. F. (2006). Patterns of Cancer Incidence, Mortality, and Prevalence Across Five Continents: Defining Priorities to Reduce Cancer Disparities in Different Geographic Regions of the World. *Journal of Clinical Oncology, 24*(14), 2308. doi:10.1200/JCO.2005.05.2308

Kandula, A. R., Sathya, R., & Narayana, S. (2023). Multivariate Analysis on Personalised Cancer Data using a Hybrid Classification Model using Voting Classifier. *International Journal of Intelligent Systems and Applications in Engineering, 11*(1), 354 - 362.

Keller, K. G., Toriola, A., & Schneider, J. K. (2021). The relationship between cancer fatalism and education. *Cancer Causes Control, 32*(2), 109 - 118. doi:10.1007/s10552-020-01363-4

Kim, H. K., & L. M. (2021). Cultural Determinants of Cancer Fatalism and Cancer Prevention Behaviours among Asians in Singapore. *Health Community, 36*(8), 940 - 949. doi:10.1080/10410236.2020.1724636

Liu, Y., Oppong, F., Adjanin, N., & Brooks, G. P. (2023). The homogeneity of Covariances Assumption in MANOVA: Differential Impact of heterogeneous Variances and

Covariances. *2023 Annual Meeting of the American Educational Research Association* (pp. 1 - 14). Philadelphia, PA: AERA.keller

Mertler, C. A., & Reinhart, R. V. (2017). *Advanced and Multivariate Statistical Methods: Practical Application and Interpretation* (Sixth ed.). New York: Taylor & Francis.

Midway, S., Robertson, M., Flinn, S., & Kaller, M. (2020). Comparing Multiple Comparisons: Practical Guidance for Choosing the Best Multiple Comparisons Test. *Peer Journal, 8*, e10387. doi:10.7717/peerj.10387

Mo, S., Dal, W., Zhou, Z., Gu, R., Li, Y., Xiang, W., . . . L, Q. (2021). Comprehensive Transcriptomic Analysis Reveals Prognostic Value of an EMT-Related Gene Signature in Colorectal Cancer. *Frontiers in Cell and Developmental Biology, 9*, 681431. doi:10.3389/fcell.2021.681431

Morgan, E., Arnold, M., Camargo, M. C., Gini, A., Kunzmann, A. T., Matsuda, T., . . . Soerjomataram, I. (2022). The current and future incidence and mortality of gastric cancer in 185 countries, 2020–40: A population-based modelling study. *EClinicalMedicine, 47*, 101404. doi:10.1016/j.eclinm.2022.101404

Mutebi, M. (2023, March 13). *Winning the Cancer War in Sub-Saharan Africa.* Retrieved July 05, 2024, from https://www.project-syndicate.org/commentary/cancer-crisis-in-africa-needs-investment-in-testing-and-prevention-by-miriam-mutebi-2023-03?utm_term=&utm_campaign=&utm_source=adwords&utm_medium=ppc&hsa_acc=1220154768&hsa_cam=12374283753&hsa_grp=117511853986&

National Cancer Institute. (2021, October 11). *What Is Cancer?* Retrieved July 05, 2024, from cancer.gov: https://www.cancer.gov/about-cancer/understanding/what-is-cancer

Olson, C. I. (1976). On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin, 83*(4), 579 - 586.

Pizzato, M., Santucci, C., Parazzini, F., Negri, E., & La Vecchia, C. (2024). Cancer Mortality Patterns in Selected Northern and Southern African Countries. *European Journal of Cancer Prevention, 33*(3), 192 - 199. doi:10.1097/CEJ.0000000000000852

Powe, B. D., & Finnie, R. (2003). Cancer Fatalism: The State of the Science. *Cancer Nursing, 26*(6), 454 - 467.

Rubio, C., Alfaro, M., Mejia-Giraldo, A., Fuertes, G., Mosquera, R., & Vargas, M. (2023). Multivariate analysis in data science for the geospatial distribution of the breast cancer mortality rate in Colombia. *Frontiers in Oncology, 12*, 1055655. doi:10.3389/fonc.2022.1055655

Salisu, W. J., Mirlashari, J., Seylani, K., Varaei, S., & Thorne, S. (2022). Fatalism, distrust, and breast cancer treatment refusal in Ghana. *Cancer Oncology Nursing Journal, 31*(2), 198 - 205. doi:10.5737/23688076322198205

Sharma, R., Nanda, M., Fronterre, C., Sewagudde, P., Ssentongo, A. E., Yenney, K., . . . Ssentongo, P. (2022). Mapping Cancer in Africa: A Comprehensive and Comparable Characterisation of 34 Cancer Types Using Estimates From GLOBOCAN 2020. *Frontiers in Public Health, 10*, 839835. doi:10.3389/fpubh.2022.839835

Smailyte, G., & Kurtinaitis, J. (2008). Cancer mortality differences among urban and rural residents in Lithuania. *BMC Public Health, 8*, 56. doi:10.1186/1471-2458-8-56

Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Boston: Pearson.

WHO. (2022, February 03). *Cancer*. Retrieved July 05, 2024, from www.who.int: https://www.who.int/news-room/fact-sheets/detail/cancer

Yanai, H., Inaba, Y., Takagi, H., & Yamamoto, S. (1979). Multivariate Analysis of Cancer Mortalities for Selected Sites in 24 Countries. *Environmental Health Perspectives, 32*, 83 - 101.