ENHANCED KNOWLEDGE IN SCIENCES AND TECHNOLOGY
e-ISSN: 2773-6385

EKST

Vol. 6 No. 1 (2026) 592-602
https://publisher.uthm.edu.my/periodicals/index.php/ekst

A Statistical Comparison of Cancer Mortality Rates Between Developed and Developing Countries

Nur Syameerah Hanani Mohd Kamaruzaman[1], Mohd Saifullah Rusiman[1*]

[1] *Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, UTHM Kampus Cawangan Pagoh, Hab Pendidikan Tinggi Pagoh, KM 1, Jalan Panchor, 84600 Pagoh, Muar, Johor, MALAYSIA.*

*Corresponding Author: saifulah@uthm.edu.my
DOI: https://doi.org/10.30880/ekst.2026.06.01.065

Article Info

Received: 4 January 2026
Accepted: 6 June 2026
Available online: 9 July 2026

Keywords

Cancer, Mortality Rates, Developed Country, Developing Country, Principal Component Analysis, Cluster Analysis, Correlation

Abstract

This study addresses disparities in cancer deaths worldwide through the analysis of long-term patterns in 20 countries in developed and developing countries by secondary data from 1990 to 2016 incorporating five major types of cancers including lung, breast, colorectal, liver, and stomach cancers. The Principal Component Analysis (PCA) is used to eliminate data dimensions and find major patterns of mortality, which proves two main components that explain a significant percentage of all the variances, the first component is significantly related to the lifestyle-associated cancer such as lung, breast, and colorectal and the second element is associated with the infection-related and region-specific mortality such as liver and stomach. According to these aspects, K-means cluster analysis (K = 2) can group the countries into high- and low-mortality groups regardless of the specified economic groups, which indicates that the statistical groupings and development status are highly consistent. The findings reveal that developed nations are largely concentrated into the category of mortality due to lung and colorectal cancer whereas developing nations are largely gathered under the category of mortality due to liver and stomach cancer showing a different epidemiology of the same arising out of the lifestyle, access to healthcare, and prevention measures. The integrative multivariate method offers strong empirical data of self-perpetuating structural disparities in the world healthcare institutions and cancer control initiatives. The current research will be useful in practice because it provides a data-based model to facilitate the implementation of specific interventions in the field of public health, a more efficient distribution of resources, and cancer prevention and early diagnosis policies that are specific to the region, which will result in the reduction of international health disparities in cancer mortality and the enhancement of international health outcomes.

1. Introduction

Global cancer mortality has remained a critical social health issue with enormous inequalities amid developed and developing nations. More developed countries have had the privilege of having better screening and medical technology and hence have been able to reduce the number of individuals that die, yet, the poor countries have a high rate of death as the diseases are diagnosed at later stages, the medical centers are not enough to meet all the demands and the resources available are insufficient [1]. The necessary gap of knowledge is bridged by the given

Enhanced Knowledge in Sciences and Technology Vol. 6 No. 1 (2026) 592-602

work because it focuses on the analytical focus of incidence data shift to the longitudinal mortality data of 20 selected countries in 1990-2016. This study quantifies the five most prevalent malignancies lung, breast, colorectal, liver and stomach cancer in terms of mortality as a last measure of the efficacy of the healthcare system which will provide an internationally consistent yardstick in gauging the global health equity [2].

The multivariate statistical procedures adopted in this study are Principal Component Analysis (PCA) and Cluster Analysis (CA) to expose the complex, latent constructs of these 27 years data. PCA is used to decrease the number of dimensions to differentiate the dominant patterns of mortality and CA is objective to cluster the nations sharing the same type of profiles regardless of their assigned economic labels [3]. Such methodologies will be able to perform an in-depth statistical comparison of developed and developing areas and determine the structure contrasts that might otherwise go unnoticed in simple descriptive statistics [4]. Eventually, the results will assist policymakers and international health institutions to create more justifiable prevention strategies and frameworks of resource distribution particularly in those environments where resources are scarce [5].

2. Research Methodology

2.1 Data Source and Description

Table 1 presents the list of variables used in this study, along with their respective descriptions and types of measurement. The table categorizes each variable as qualitative or quantitative and specifies whether the data are categorical or continuous, providing a clear overview of the dataset structure and the role of each variable in the analysis of cancer mortality across countries and years.

Table 1 List of Variable

No	Variables	Description	Types of Variables
1	Country	Country name for cancer death	Qualitative and categorical variables
2	Year	Represents cancer death per year	Quantitative and continuous variable
3	Population	Population of people in each country by year	Quantitative and continuous variable
4	Liver cancer	Cancer name	Quantitative and continuous variable
5	Breast cancer	Cancer name	Quantitative and continuous variable
6	Stomach cancer	Cancer name	Quantitative and continuous variable
7	Colon and rectum cancer	Cancer name	Quantitative and continuous variable
8	Tracheal, bronchus, and lung cancer	Cancer name	Quantitative and continuous variable

The data used are the mortality rates of cancer per annum (1990 to 2016). Twenty countries were sampled which included 10 developed countries (United States, United Kingdom, Germany, France, Canada, Japan, Australia, Sweden, Netherlands, and South Korea) and 10 developing countries (Bangladesh, India, Pakistan, Indonesia, Philippines, Nepal, Nigeria, Kenya, Ethiopia and Tanzania) [6]. There were five predominant types of cancer that were chosen because of their global occurrence and the number of cancer-related deaths they cause: lung cancer, breast cancer, colorectal cancer, liver cancer, and stomach cancer [7]. The mortality rates were taken as the number of deaths per 100,000 population. The age-standardized mortality rate (deaths per 100,000 population) is the main outcome variable that will remove the demographic bias introduced by the variations in the age structure of the global population. The study article also gives a consistent and impartial comparison of world cancer mortality by eliminating other external variables such as health expenditure or lifestyle. The cancer mortality rate was calculated using the formula in Equation (1) [8]:

$$Cancer\ Mortality\ Rate = \left(\frac{Number\ of\ Cancer\ Deaths}{Population}\right) \times 100,000 \quad (1)$$

2.1.1 Data Pre-Processing

Data pre-processing starts by combining a cancer death record with a corresponding population record by country and year, then resulting in a mortality rate to balance the data with the population size as opposed to absolute death rates. Missing values were addressed through external official means, or eliminated to represent the dataset integrity, with the aim at having a stable 27-year longitudinal dataset. Standardization was performed using the following formula in Equation (2) [9]:

$$Z = \frac{X - \mu}{\sigma} \quad (2)$$

where Z is the standard form of a variable, X is the original value of the variable, μ is the mean of the variable and 0.79 is the standard deviation. Standardization of the data is a very vital step to Principal Component Analysis because it enables every variable to contribute equally to the end components.

2.2 Methodology

2.2.1 Principal Component Analysis

The Principal Component Analysis (PCA) is a dimensionality reducing tool that is used to simplify the 27-year data set by converting correlated mortality rates into a few uncorrelated factors known as principal component in Equation (3) [10]:

$$Z_t = \alpha_{i1}X_1 + \alpha_{i2}X_2 + \cdots + \alpha_{in}X_n \quad (3)$$

Equation 3 shows the equation of first Z-score standardizing the data, hence ensuring that all forms of cancer have an equal contribution and then analysis of the eigenvalues, to identify such major causes of mortality signatures, which may be an infection-related cluster in the developing countries or life-style patterns which are widespread in the developed countries [11]. The study adopts an objective approach by grouping countries based on shared healthcare challenges and epidemiological trends rather than predefined economic categories, using the extracted component scores as the basis for subsequent cluster analysis. This integrated approach assists in simplifying the complexity of a dataset of 27 years to stable and decodable dimensions that provides a powerful empirical background to determine global health inequities and determine a certain method of prevention-focused health policies [12].

2.2.2 Cluster Analysis

Cluster Analysis (CA) is used to cluster 20 countries according to their similarities in their mortality rates of five major cancers without using the economic labels that are previously set in advance to explore the natural health trends. The study adopted the algorithm of K-means clustering whereby K=2 is used to test the hypothesis of dichotomous structure between high and low mortality profiles [13]. This is done by standardizing the mortality rates in Z-scores to make the searches equal-weighted then repeated assignment of each country to the nearest cluster centroid by minimizing the Within-Cluster Sum of Squares (WCSS) as expressed by the objective formulation as in Equation (4) [14]:

$$min \sum_{k=1}^{K} \sum_{x_i \epsilon s_k} \|x_i - \mu_k\|^2 \quad (4)$$

These clusters obtained empirically (Cluster 1 and Cluster 2) are then contrasted with traditional socioeconomic categories to identify whether structural variations in cancer mortality truly represent global differences in healthcare facilities, early detection and management of diseases.

3. Results and Discussion

3.1 Overall Trends in Cancer Mortality Rates

This part investigates global cancer mortality rates between 1990 and 2016 on liver, breast, stomach, colorectal, and lung cancer. Employing descriptive statistics to determine central tendencies and dispersion, the study determines the main contributor to the death burden of malignancies in the world and how they vary with time. This introductory examination gives the needed background to further multivariate techniques, and this initial

examination gives a preliminary overview of the mortality differences between developed and developing countries.

3.1.1 Descriptive Statistics of Cancer Mortality Rates

Table 2 presents the descriptive statistics for each cancer type, including the mean, minimum, maximum, and standard deviation of mortality rates, providing an overview of the central tendency and variability across the study period.

Table 2 Descriptive Statistics for Each Types of Cancer

Variables	Mean	Minimum	Maximum	Standard Deviation
Liver Cancer Mortality Rates	6.7630	1.03781	31.8089	6.9498
Breast Cancer Mortality Rates	10.5958	1.6400	27.9426	7.7257
Stomach Cancer Mortality Rates	10.5010	1.1058	43.0854	9.9349
Colon and Rectum Cancer Mortality Rates	15.9988	1.7421	39.0182	13.5156
Tracheal, Bronchus, and lung Cancer Mortality Rates	26.8200	1.2115	70.6523	23.8832

This part provides descriptive statistics of the mortality rates of five major malignancies lung, colorectal, breast, stomach and liver cancer standardized per 100,000 individuals in 20 countries in 1990 to 2016. Lung cancer is the health burden with the highest average mortality rate (26.82) and variability (SD = 23.88), second is colorectal (15.99), breast (10.59), and stomach cancer (10.50), and liver cancer had the lowest average (6.76). The high values of standard deviation, which are high between minimum and maximum values, and the significant values between minimum and maximum values, in the lung and stomach cancers, indicate the existence of significant geographical and temporal variations. Such differences indicate that local risk factors, inequity in access to healthcare, and localized success of prevention strategies dominate the outcome of mortality.

3.1.2 Trend Analysis Over Time (1990-2016)

Fig. 1 illustrates the 1990 to 2016 trend analysis indicated a landscape diversified scenario of the global cancer mortality with the main cause being the steady and high increasing trend of lung cancer.

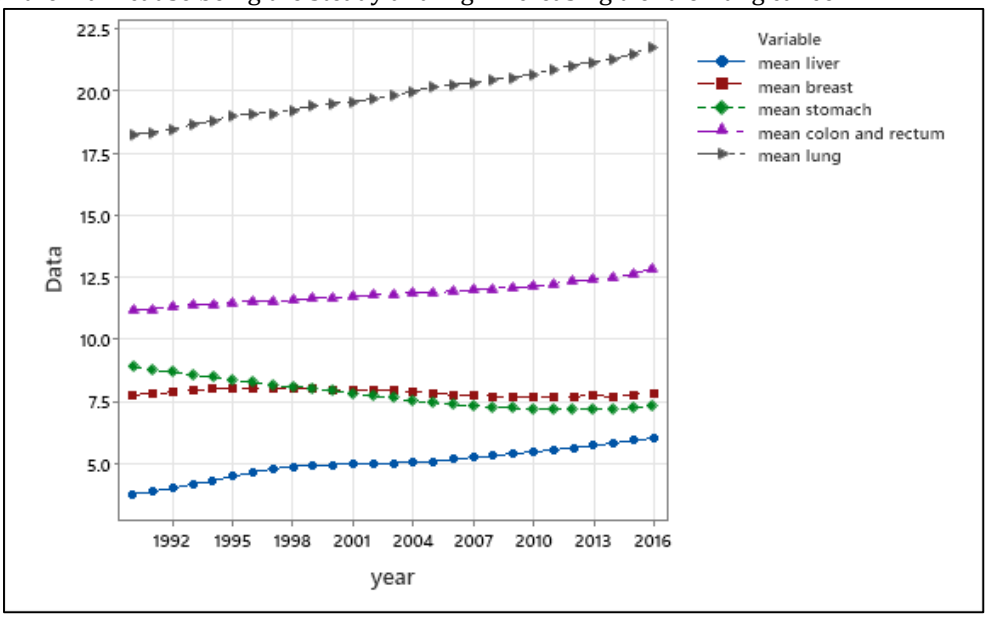

Fig. 1 Trend Analysis Over Time

Conversely, stomach cancer is the only cancer that has shown a consistent long-term decrease, and it will eventually surpass the death rates of breast cancer by the middle of the 2000s. Although liver and colorectal cancer have shown to grow at different rates, breast cancer is mortality cancer that has the least trend over the

27-year interval, with the smallest slope. All these trends cumulatively point to the fact that, although most of the leading cancers are beginning to experience rising mortality rates, the burden of stomach cancer is being effectively reduced through better intervention or altered risk factors.

3.1.3 Comparison Between Cancer Types

Table 3 compares average cancer mortality rates in 1990 and 2016, highlighting trend patterns and changes across different cancer types over the study period.

Table 3 Comparison of Average Cancer Mortality Rates Between 1990 and 2016

Cancer Type	Trend Pattern	1990 Rate	2016 Rate
Tracheal, Bronchus, and Lung Cancer	Increasing	18.22	21.73
Colon and Rectum Cancer	Increasing	11.16	12.83
Stomach Cancer	Decreasing	8.90	7.35
Breast Cancer	Stable	7.76	7.81
Liver Cancer	Increasing	3.76	6.02

The 1990-2016 trend analysis indicates that the mortality increase was the most significant and constant in the case of lung cancer. Stomach cancer is the only one to show a linear long-term decrease, however, implying that preventive measures and control of risk factors are improving. There are also increasing trends in liver and colorectal cancers though at varying rates. Breast cancer does not vary greatly during the period, and it is the least changed. Overall, these findings suggest that though most of the major cancers present an increasing mortality rate, the weight of stomach cancer has been decreasing over the years.

3.2 Principal Component Analysis

3.2.1 Scree Plot of PCA

Fig. 2 presents the scree plot of eigenvalues from the principal component analysis, illustrating the contribution of each component to the total variance in the cancer mortality variables.

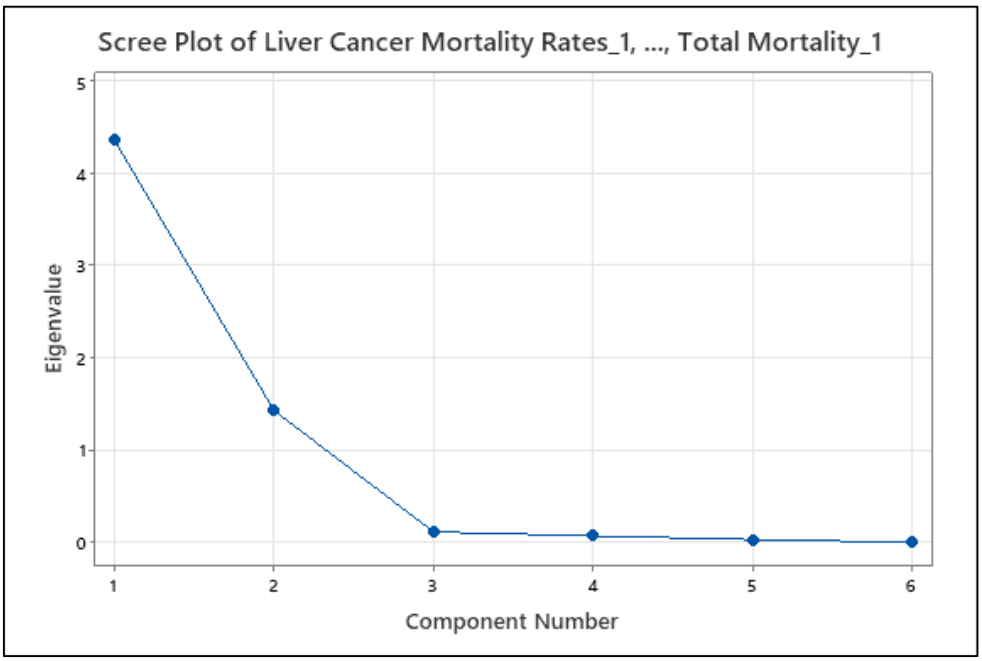

Fig. 2 Scree Plot for All Cancer

The graph indicates that the eigenvalues decrease drastically between the first and the second principal component and then there is a definite elbow and a slow-going smooth upward inclination in the subsequent components. This trend shows the initial two components describe much of the overall variance and represent the most significant structure of the cancer mortality data and the remaining components add comparatively less

and primarily exhibit smaller variations or noise. Thus, the two main elements can offer a sensible and efficient model of the initial multidimensional data and enable easier recognition of major trends in mortality and make comparisons between developed and developing areas possible.

3.2.2 Matrix Plot for All Cancer Types

Fig. 3 shows a matrix plot of pairwise relationships among liver, breast, stomach, colon and rectum, and lung cancers, with Pearson correlation coefficients and 95% confidence intervals.

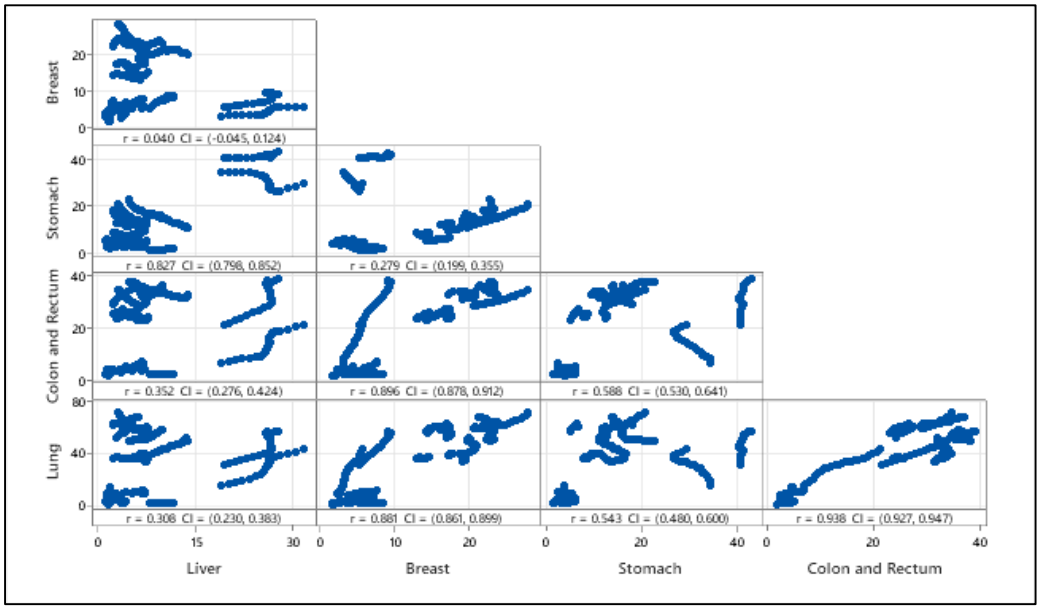

Fig. 3 Matrix Plot for All Cancer

The Pearson correlation matrix indicates that the five types of cancers show significant positive correlations with one another, thus the most appropriate statistical justification of employing Principal Component Analysis (PCA) to overcome the problem of data multicollinearity. Most pronounced correlations can be observed between the breast, colorectal and lung cancer, with the latter two being the most significant ones (lung and colorectal have a correlation of about (r=0.94). Although liver cancer has fewer sensitive relationships with breast cancer (r = 0.04), it has moderate relationships with stomach and colorectal cancer (r = 0.53 and r = 0.35 respectively). These small confidence intervals show that the observed redundancies are statistically significant, and hence the presence of a dimension-reduction method such as PCA is necessary to compress these overlapping variables to few non-correlated, dominant mortality patterns.

3.2.3 Preliminary Correlation Analysis of Cancer Mortality Rates

Table 4 correlation analysis indicates that most of the cancer mortality rates are significantly and positively correlated, which meets the requirement of Principal Component Analysis (PCA).

Table 4 Correlation Matrix of Cancer Mortality Rates

	Cancer Type	Liver Cancer	Breast Cancer	Stomach Cancer	Colorectal Cancer	Lung Cancer
Correlation	Liver Cancer	1.000	0.040	0.827	0.352	0.308
	Breast Cancer	0.040	1.000	0.279	0.896	0.881
	Stomach Cancer	0.827	0.279	1.000	0.588	0.543
	Colorectal Cancer	0.352	0.896	0.588	1.000	0.938
	Lung Cancer	0.308	0.881	0.543	0.938	1.000

Sig. (1-tailed)	Liver Cancer		0.177	< 0.001	< 0.001	< 0.001
	Breast Cancer	0.177		0.000	0.000	0.000
	Stomach Cancer	0.000	0.000		0.000	0.000
	Colorectal Cancer	0.000	0.000	0.000		0.000
	Lung Cancer	0.000	0.000	0.000	0.000	

Although there is no meaningful correlation between liver and breast cancer (there is no significant result, and the value of 0.040 is insignificant, and the correlation is 0.177), liver cancer is significantly related to stomach and lung cancer, and colorectal cancer is almost perfectly correlated with lung cancer ($r = 0.938$, $p = 0.001$). The results of these extensive, statistically significant correlations suggest a very high level of multicollinearity and redundant data in the dataset. Therefore, it can be claimed that using PCA is reasonable to reduce these confounding variables to a smaller number of uncorrelated factors that will make it easier to identify the primary factors behind the differences in cancer mortality in the world.

3.2.4 Principal Component Biplot Interpretation

Fig. 4 displays a PCA biplot illustrating the relationships between countries and cancer mortality variables based on the first two principal components, with vectors indicating the contribution and direction of each cancer type.

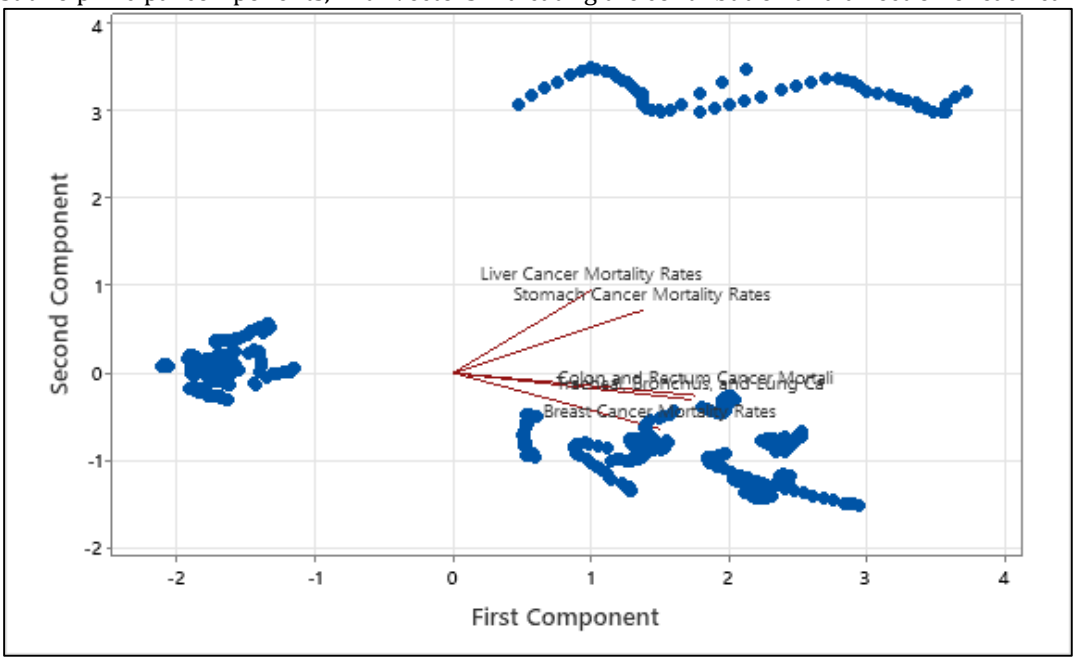

Fig. 4 Biplot of All Cancer Types

The PCA Biplot in Fig. 4 gives a complete picture by superimposing the country-year observation and the vectors of cancer mortality loading display how local cancerous conditions characterize the mortality trend by region. As the analysis shows, there is a particular spatial distribution with the developed Western countries such as the United States, the United Kingdom, and Australia being around the Breast and Colorectal cancer vectors which are associated with mortality burdens related to Western lifestyles. On the other hand, Japan and South Korea are the dominant vectors of Stomach and Liver cancer, which points to a rather different epidemiological picture as high hepatic and gastrointestinal burden are not typical of the West despite the high level of development. This biplot makes good sense to understand the way that systematic inequalities in healthcare and environment translate into discrete mortality imprints in various parts of the world.

3.3 Cluster Analysis Results

3.3.1 Selection Number of Cluster (K=2)

Table 5 shows the distance between the final cluster centers from the K-Means clustering analysis for each cancer type. The distance values indicate how well the clusters are separated, with higher values representing clearer separation between clusters.

Table 5 Distance Between Final Cluster Centers for K-Means Clustering

Cancer Type	Distance Between Cluster	Interpretation
Liver Cancer	20.955	Clear separation between clusters
Breast Cancer	14.692	Moderate but meaningful separation
Stomach Cancer	27.776	Strong separation between clusters
Colon and Rectum Cancer	26.016	Strong separation between clusters
Tracheal, Bronchus and Lung Cancer	44.846	Very strong separation between clusters

Discussing the analysis of the last cluster in Table 5, it can be stated that there are clear distinctions between all the five cancer types that confirm that K = 2 should be used to determine two distinct patterns of mortality. Lung cancer records the highest distance (44.846), which means that there are major disparities that probably are caused by smoking rate, industrialization and environmental policy. There is also a significant difference of stomach, colorectal and breast cancer indicating the differences in diet, urbanization and early screening and treatment between the developed and developing nations. In general, these findings support the fact that the trends in cancer mortality are not homogenous all over the world and can be clustered into two distinctly different groups.

3.3.2 K-Means Clustering Results

Table 6 presents the final cluster centers and ANOVA results from the K-Means clustering analysis for each cancer type. The table shows the mean values for Cluster 1 and Cluster 2, along with the F-values and p-values, which indicate the statistical significance of differences between the two clusters.

Table 6 Final Cluster Center and ANOVA Results for K-Means Clustering

Cancer Types	Final Cluster 1	Final Cluster 2	F-value	p-value
Liver Cancer	4.6675	25.6224	2446.677	< 0.001
Breast Cancer	4.7190	19.4110	3586.576	< 0.001
Stomach Cancer	43.0854	1.1058	1438.995	< 0.001
Colon and Rectum Cancer	3.9060	29.9224	6490.103	< 0.001
Tracheal, Bronchus, and Lung Cancer	49.9903	5.1446	4025.377	< 0.001

Each of the clusters and the results and findings of ANOVA validates the apparent statistical difference between two clusters: Cluster 1, which is characterized by relatively low mortality picture and Cluster 2, which is characterized by a systematically higher rate of mortality in all the five cancer types. The distance between the two centers, particularly the stomach and lung cancers, is quite large which accentuates brutal disparities in disease prevalence on an international level which can be explained by the systemic differences in care accessibility and diagnostic equipment. Besides, F-values and the p-values are high, and the ANOVA test shows that each type of cancer is contributing significantly to this set of groupings, and, therefore, the groupings that have been formed are not arbitrary and of a high level of statistical strength. This quantitative basis provides the justification that the K-means algorithm worked to categorize the countries into two distinct mortality profiles that provides a good basis on which the discussion can be made between developed and developing nations.

3.3.3 Cluster Membership of Countries

Table 7 shows the cluster membership of selected countries based on cancer mortality rates for five cancer types. Each country is classified into either Cluster 1 (C1) or Cluster 2 (C2), allowing for comparison of cancer mortality patterns across countries with different levels of development.

Table 7 Cluster Membership of Countries Based on Cancer Mortality Rates

Country	Liver	Breast	Stomach	Colorectal	Lung
Australia	C1	C2	C2	C2	C1
Canada	C1	C2	C2	C2	C1
France	C1	C2	C2	C2	C1
Germany	C1	C2	C2	C2	C1
United Kingdom	C1	C2	C2	C2	C1
United State	C1	C2	C2	C2	C1
Japan	C2	C1	C1	C2	C1
Netherland	C1	C2	C2	C2	C1
South Korea	C2	C1	C1	C1	C1
Sweden	C1	C2	C2	C2	C1
Bangladesh	C1	C1	C2	C1	C2
Ethopia	C1	C1	C2	C1	C2
India	C1	C1	C2	C1	C2
Indonesia	C1	C1	C2	C1	C2
Nepal	C1	C1	C2	C1	C2
Nigeria	C1	C1	C2	C1	C2
Kenya	C1	C1	C2	C1	C2
Pakistan	C1	C1	C2	C1	C2
Philippine	C1	C1	C2	C1	C2
Tanzania	C1	C1	C2	C1	C2

The cluster membership analysis indicates that the status of national development is a major factor that determines cancer mortality, but there are also regional and epidemiological trends in Table 7. The economic line is strict in colorectal cancer since all developed countries are in the high-mortality group, which is probably related to Western diets and aging. On the other hand, the deaths of lung cancer are maximum in all the developing countries and South Korea, which demonstrates high smoking rates and pollution in these countries. Curiously, the liver and stomach cancer status isolates Japan and South Korea; they share the highest liver cancer mortality with lower mortality rates in stomach cancer, which developed high-quality national screening in comparison with the rest of the world. Breast cancer has been a major burden of Western industrialization with the highest mortality in eight western developed countries. In general, recurrent aggregation of developing nations into high-mortality categories in various cancers demonstrates a high level of system-wide inequality in medical achievement in the world.

4. Conclusion and Recommendations

This study presents that cancer mortality rates in the world between the years 1990 and 2016 are highly dependent on the level of a country's development and access to medical facilities. By using Principal Component Analysis, the intricate connections among the five leading types of cancer were limited to a few major components that accounted for most of the variations in cancer mortality around the world. These results were also supported by the K-means cluster analysis, which clustered the countries into the high and low-mortality groups. The consequent clusters are mostly data-driven, and reflect on the long-term socioeconomic disparities, emphasizing the idea that the disparities in the screening, diagnosis, and the success of treatment are the significant contributors to the global cancer-related outcomes.

In order to overcome these global health inequalities, policymakers in developing nations ought to focus on the need to invest in early diagnosis and standardized screening programmed especially on high-burden cancers, such as lung, liver and stomach cancer. Future studies may also extend this study by incorporating more variables,

like healthcare spending, personal risk factors, and socioeconomic measures to give a more comprehensive view of what is causing these clusters. Besides, the application of sophisticated longitudinal approaches, such as time-series analysis or hierarchical clustering with post 2016-based data, would serve to keep findings up to date. Lastly, the weaknesses of cancer registry systems in many countries should be reinforced to enhance accuracy of data and facilitate effective and evidence-based interventions to minimize preventable deaths related to cancer in the world.

Acknowledgement

The author would like to thank the Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, for its support.

Conflict of Interest

Authors declare that there is no conflict of interest regarding the publication of the paper.

Author Contribution

The authors confirm that the contribution to the article is as follows: ***study conception and design****: Nur Syameerah Hanani Mohd Kamaruzaman, Mohd Saifullah Rusiman;* ***data collection****: Nur Syameerah Hanani Mohd Kamaruzaman;* ***data analysis and interpretation of results:*** *Nur Syameerah Hanani Mohd Kamaruzaman, Mohd Saifullah Rusiman;* ***draft manuscript preparation****: Nur Syameerah Hanani Mohd Kamaruzaman, Mohd Saifullah Rusiman. All authors read and approved of the final manuscript.*

References

[1] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA a Cancer Journal for Clinicians, 71(3), 209–249. https://doi.org/10.3322/caac.21660

[2] Peravali, M., Joshi, I., Ahn, J., & Kim, C. (2021). A Systematic Review and MetaAnalysis of Clinical Characteristics and Outcomes in Patients With Lung Cancer with Coronavirus Disease 2019. JTO Clinical and Research Reports, 2(3), 100141. https://doi.org/10.1016/j.jtocrr.2020.100141

[3] Mao, Y., Gao, Y., He, Y., Wan, Z., Li, S., Ding, Z., Hu, B., Fu, L., Luo, C., Zhu, S., & Cao, W. (2025). Global burden of cancer in women, 1990–2021: a systematic analysis from the GBD 2021 study. Frontiers in Oncology, 15, 1633894. https://doi.org/10.3389/fonc.2025.1633894

[4] Walker, N., Heuer, A., Sanders, R., & Tong, H. (2024). The costs and benefits of scaling up interventions to prevent poor birth outcomes in low-income and middle-income countries: a modelling study. The Lancet Global Health, 12(9), e1526–e1533. https://doi.org/10.1016/s2214-109x(24)00238-9

[5] Lin, S., Gao, K., Gu, S., You, L., Qian, S., Tang, M., Wang, J., Chen, K., & Jin, M. (2021). Worldwide trends in cervical cancer incidence and mortality, with predictions for the next 15 years. Cancer, 127(21), 4030–4039. https://doi.org/10.1002/cncr.33795

[6] Gultekin, M., Karaca, M. Z., Kucukyildiz, I., Dundar, S., Keskinkilic, B., & Turkyilmaz, M. (2019). Mega Hpv laboratories for cervical cancer control: Challenges and recommendations from a case study of Turkey. Papillomavirus Research, 7, 118–122. https://doi.org/10.1016/j.pvr.2019.03.002

[7] Huang, J., Ssentongo, P., & Sharma, R. (2023). Editorial: Cancer burden, prevention and treatment in developing countries. Frontiers in Public Health, 10. https://doi.org/10.3389/fpubh.2022.1124473

[8] Ma, J., Jemal, A., Fedewa, S. A., Islami, F., Lichtenfeld, J. L., Wender, R. C., Cullen, K. J., & Brawley, O. W. (2019). The American Cancer Society 2035 challenge goal on cancer mortality reduction. CA a Cancer Journal for Clinicians, 69(5), 351–362. https://doi.org/10.3322/caac.21564

[9] Tuppad, N. A., & Patil, N. S. D. (2023). Data pre-processing issues in medical data classification. Journal of Advanced Zoology, 44(S6), 1079–1084. https://doi.org/10.17762/jaz.v44is6.2361

[10] Tang, T. M., & Allen, G., I. (2018). Integrated Principal Components analysis. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1810.00832

[11] Salachan, P. V., Rasmussen, M., Fredsøe, J., Ulhøi, B., Borre, M., & Sørensen, K. D. (2022). Microbiota of the prostate tumor environment investigated by wholetranscriptome profiling. Genome Medicine, 14(1). https://doi.org/10.1186/s13073-022-01011-3

[12] Borges, V. R. P., Esteves, S., De Nardi Araújo, P., De Oliveira, L. C., & Holanda, M. (2018). Using Principal Component Analysis to support students' performance prediction and data analysis. Anais Do . . . Simpósio Brasileiro De Informática Na Educação/Anais Do Simpósio Brasileiro De Informática Na Educação, 1, 1383. https://doi.org/10.5753/cbie.sbie.2018.1383

[13] Yilmaz, S., Janelsins, M., Flannery, M., Culakova, E., Wells, M., Lin, P., Loh, K., Epstein, R., Kamen, C., Kleckner, A., Norton, S., Plumb, S., Alberti, S., Doyle, K., Porto, M., Weber, M., Dukelow, N., Magnuson, A., Kehoe, L., . . .

Mohile, S. (2022). Protocol paper: Multi-site, cluster-randomized clinical trial for optimizing functional outcomes of older cancer survivors after chemotherapy. Journal of Geriatric Oncology, 13(6), 892–903. https://doi.org/10.1016/j.jgo.2022.03.001

[14] Cuello, M. A., Gomez-Valenzuela, F., Wichmann, I., & Olawaiye, A. B. (2025). Global determinants of gynecologic cancer incidence and mortality: A cluster- based analysis with predictive insights. International Journal of Gynecology & Obstetrics, 171(S1), 147–165. https://doi.org/10.1002/ijgo.70279