ISSN: 2996-2625

Research Article

Medicine & Pharmacology: Open Access

Clusters of Communicable Diseases in Selected Regions of Tanzania

Twalibu Fadhili and Ramkumar T. Balan*

CNMS, University of Dodoma, Tanzania

***Corresponding Author**
Ramkumar T. Balan, CNMS, University of Dodoma, Tanzania

Submitted: 2025, May 19; **Accepted:** 2025, Jun 25; **Published:** 2025, Jul 01

Citation: Fadhili, T., Balan, R. T. (2025). Clusters of Communicable Diseases in Selected Regions of Tanzania. *Med Pharmacol OA, 3*(2), 01-11.

Abstract

All around the world, particularly in Africa, communicable diseases have become a challenge to the public health system. Understanding the interactions and dynamics between multiple diseases that share resources, time, and geography is essential, especially in contexts where resources are scarce. This study is the beginning in Tanzania to assess the communicable disease clusters in selected regions to make precautions and awareness amongst people and the healthcare community to prevent the diseases. Four diseases—influenza, urinary tract infection, diarrhoea, and malaria, depending on their mode of transmission—vector, person-to-person contact, oral contact, and airborne contact, respectively—were considered for the study. The study includes six zones: Arusha, Dar es Salaam, Dodoma, Mbeya, Mtwara, and Mwanza for the period of 2020 to 2021. Statistical tools like discriminant analysis and cluster analysis pinpointed the risky age groups with multiple diseases in specified regions. The study showed that the elderly in Arusha, Mbeya, and Mtwara were at high risk, while the children in Dar es Salaam and Dodoma and the infants in Mwanza were also at risk.

Keywords: Communicable Disease, Discriminant Analysis, Clusters, High Risk, Four Diseases Influenza, Urinary Tract Infection, Diarrhoea, Malaria, Transmission Air Borne, Head To Head, Oral, Vector Borne, Regions Arusha, Dar Es Salaam, Dodoma, Mbeya, Mtwara, Mwanza, Risk Levels, Low Medium and High Communicable Disease, Discriminant Analysis, Mahalnobis' Distance, Cluster Analysis, Agglomerative Method

1. Introduction

Communicable diseases (DCs) are more widely acknowledged as a serious worldwide public health issue that has a negative impact on people's health, healthcare expenses, and productivity [1-3]. Despite there being enormous advancements in prevention and treatment, CDs continue to be the world's major causes of death, morbidity, and deterioration of living conditions for millions of people [4-7]. However, Africa faces challenges with public health management due to lack of disease knowledge, prevalence and awareness, poor information, and inadequate health resources such as specialists, infrastructure, medical equipment, and medicine [8]. Thus, by considering the prevalence of multiple diseases, categorizing disease occurrences, and clustering disease-prone zones, it is necessary to control the diseases effectively. It will help to determine the extent, variations, and consequences of the chosen diseases based on their modes of transmission, such as airborne, oral, vector-borne, and head-to-head in different locations reveal three means of transmission oral, vector-borne, and airborne and include salmonellosis, borreliosis, and influenza, respectively, in the Arctic and Subarctic regions [9,10]. The study's findings show that there is a strong regression between climatic covariates and diseases caused by contaminated food and water. Moreover,conducted research on three communicable diseases, including the Human Immunodeficiency Virus (HIV), Tuberculosis (TB), and Hepatitis C Virus (HCV), amongst prisoners and the public. The study's findings show that the reduction of imprisonment rates was predicted to lower TB infections amongst inmates as well as HIV and HCV infections amongst drug injectors [11].

Considering the literature, univariate studies were more frequently found to estimate the prevalence and the spread of the disease, while the rate of incidence of disease was analysed using the multivariate strategy. Techniques on several variables will produce clinical judgments for the study that are more accurate and trustworthy

[12-15]. This study used a multivariate analysis to classify four diseases—diarrhoea, influenza, malaria, and urinary tract infection (UTI) in six locations as frequent incidence amongst different age groups. Discriminant analysis and cluster analysis were used to categorize the diseases in selected regions at low, moderate, and high risk. The statistical techniques reviewed for CDs are based on various strategies related to clusters of communicable diseases in particular places that have been categorized or identified. According to Gupta et al. (2022), Linear Discriminant Analysis (LDA) is the component of classification that optimizes the separation of the data, which maximizes the difference between the averages and minimizes the variance in accounting for most categories. Rasheed (2022) investigated extensively modified reverse transcriptase-polymerase chain reaction in the COVID-19 pandemic by choosing the optimal linear discriminants that maximize variability across numerous classes. He suggested that categorizing performance by assessing the influence of clinical data should be taken into consideration under LDA for prediction purposes.

In Tanzania, the use of linear discriminant analysis to estimate prevalence and incidence rate is not yet explored. Since the goal of the study was to allocate scarce resources and to display the disease categories of the multiple diseases in the chosen locations in their corresponding zones, LDA was the best method to get accurate estimations of the course of the disease as well as the allocation of Tanzania's limited resources [16]. This was a cross-sectional study using secondary data from the Ministry of Health (MoH) database and was based on four diseases, including diarrhoea, influenza, malaria, and UTI. It was conducted at the Monitoring and Evaluation Department (MED), which is based on the Tanzanian mainland, gathered and organized monthly from 2020 to 2021. A variety of unity types zones, and regions were used in the study, including the northern zone Arusha, the coastal zone Dar es Salaam, the central zone Dodoma, the southern highland Mbeya, the southern zone Mtwara, and the Lake Zone Mwanza. All patients recorded in the MoH database, regardless of age groups, were included in the study population.

2. Methodology

Using a linear discriminant analysis, the risk at different age groups was determined, where the sample unit included incidences of diseases diarrhoea, influenza, malaria, and UTI and patient occurrences infant, childhood, adolescence, and elders as covariates. The objective of categorizing the disease affected as minor, moderate, and high was done by discriminant functions. Group means, coefficients of lineardiscriminants, and prior probabilities of groups were taken into account throughout the analysis to distinguish the zones with respect to diseases, where the coefficients for the linear discriminant function are considered [17,18]. Wichiern (2007) reveals that Mahalanobis and prior probability are used to create a linear discriminant function, and the group that has the shortest Mahalanobis distance in terms of multivariate space is the group to which it is most similar. Moreover, the observations of a case are presented to the linear function of each group in order to get a classification score; the case is then allocated to the group with the highest classification score. By applying the Bayes theorem to these categorization scores, posterior probabilities are calculated.

3. Results and Discussions

Using a linear discriminant analysis, the risk in different age groups was determined, where the sample unit included incidences of diseases diarrhoea, influenza, malaria, and UTI and patient occurrences infant, childhood, adolescence, and elderly as covariates. The objective of categorizing the disease affected as minor, moderate, and high was done by discriminant functions. Group means, coefficients of linear discriminants, and prior probabilities of groups were taken into account throughout the analysis to distinguish the zones with respect to diseases. Wichiern (2007) reveals that shortest Mahalanobis distance and prior probability are used to create a linear discriminant function, and similar and dissimilar groups. Moreover, the observations of a case are presented to the linear function of each group in order to get a classification score; the case is then allocated to the group with the highest classification score. By using the Bayes theorem for these categorization scores, posterior probabilities were calculated; that is, the assigned categorization is assessed using posterior probabilities of members of the groups and typicality probabilities, which are measures of how typical a case's data is for each category [19].

3.1. Results

The linear discriminant functions were applied to the incidence data set for the four diseases, which included the total number of cases throughout the study areas. The study focuses on the spatial distribution of the chosen diseases by categorizing the age groups into low, medium, and high risks. The results of selected regions were as follows:

3.2. Arusha Region

There are three discriminant functions generated after grouping data on the measured covariates. The mathematical formulation of the discriminant function is provided below

$$Z_1 = -0.0013489\,A - 0.0004770926B + 0.00132044C + 0.0019802474D \qquad 1$$

$$Z_2 = 0.0145472\,A - 0.0014256251B - 0.0002573747C + 0.001251644D \qquad 2$$

$$\mathrm{Z}_3 = -0.002241038\mathrm{A} + 0.001213197\mathrm{B} - 0.00124355\mathrm{C} + 0.00638875\mathrm{D} \qquad 3$$

A = infant (age < 1 year), B =childhood (1≤ age < 5 year), C = adolescence (5 year ≤ age < 60 years) and D =elders (age ≥ 60).

The first discriminant function, Z1, showed that D had the largest contribution to differentiate the four diseases. The scale rates that signify covariates were obtained as follows: Z1 captured 92.96% of the differences between the diseases (D, A, C, B). The second and third discriminant functions, Z2 and Z3, together explained only 7.025% and 0.03% of the total variance, respectively. Also, from Figure 1, the UTI was positively influenced at high concentrations by covariates C and D, whereas diarrhoea, malaria, and influenza are unfavourable at high concentrations to covariates A and B. That is, in Arusha, children over the age of five suffer significantly from UTI, while children under the age of five suffer significantly from malaria, diarrhoea, and influenza. The results showed that 91.7% of the samples taken from various diseases were accurately matched to the associated covariates, while 8.3% of them had misclassification rates.

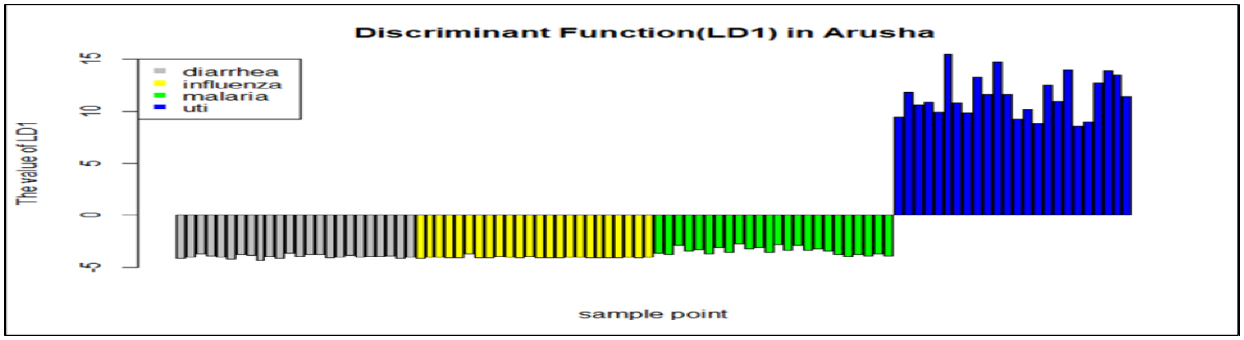

Figure 1: Displaying the results of the first discriminant function for disease incidences in Arusha.

3.3. Dar es Salaam Region

The mathematical formulation of the discriminant functions is as follows: these two functions depend on their posterior probabilities as follows.

$$Z_1 = 0.062A + 0.101B + 0.025C + 0.064D \quad 4$$

$$Z_2 = 0.001A + 0.0008B - 0.00035C + 0.00065D \quad 5$$

according to equation Z1, B contributed the most to differentiate the four diseases. Z1 predicted 98.8% of the variations in the diseases and was ranked B, D, A, and C in order of significance. Z1 has outdone Z2 in capturing 1.2% of the overall variance. Moreover, UTI has a favourable impact on all covariates, as shown by the study's result in Figure 2. Results show that UTI is the major disease of people in Dar es Salaam of all ages. The study demonstrates that 100% of the samples collected from the study were accurately assigned to their relative covariates with respect to LDA.

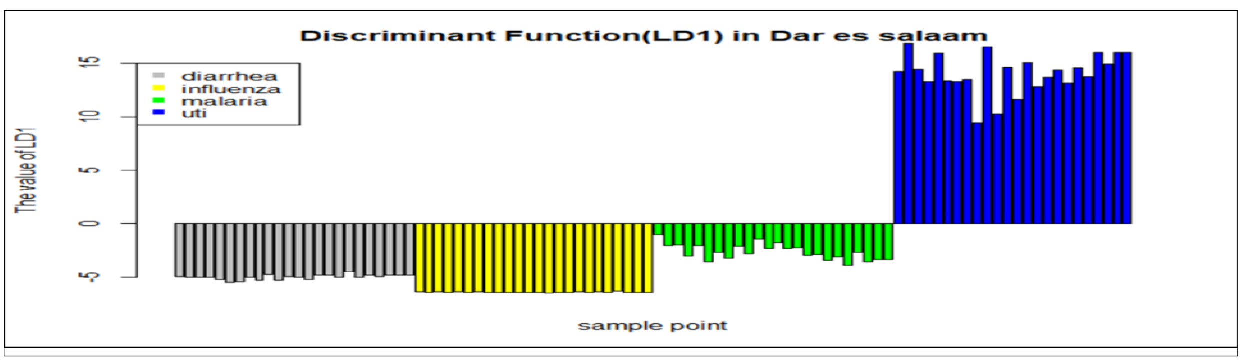

Figure 2: Displaying the first discriminant function for disease incidences in Dar es Salaam.

3.3. Dodoma Region

The discriminant functions are

$$Z_1 = 0.00058A - 0.00088B + 0.00081C + 0.00076D \quad 6$$

$$Z_2 =- 0.149A + 0.112B - 0.0645C + 0.026D \quad 7$$

$$Z_3 =- 0.007A + 0.0039B - 0.00056C - 0.00142D \quad 8$$

Z1 reveals that B had the greatest influence on the ability to distinguish between the four diseases. The discriminant functions Z2 and Z3 contributed only 10.66% and 0.07% of the total variance respectively. As shown in Figure 3, UTI had a positive influence on A, C, and D, whereas influenza has a negative effect on B. Moreover, in the study, 91.7% of the samples studied were correctly categorized into their relevant characteristics, whereas 8.3% of the samples were incorrectly classified.

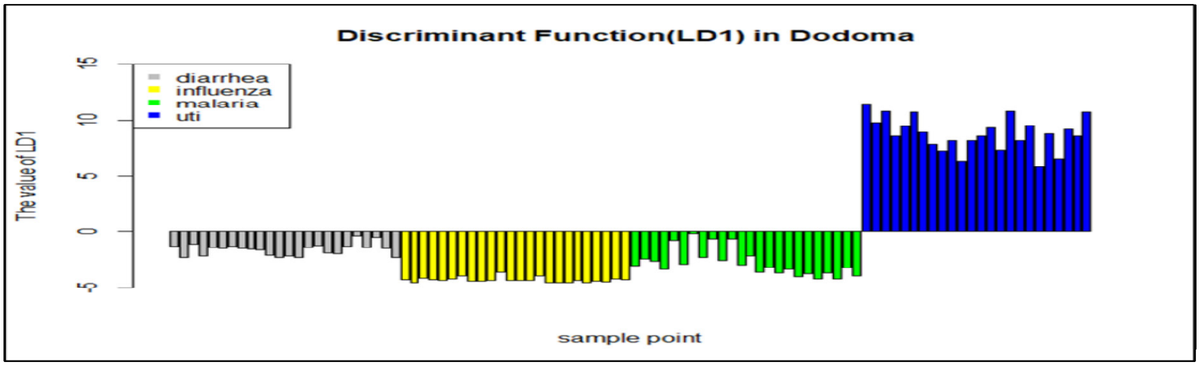

Figure 3: Displaying the results of the first discriminant function for disease incidences in Dodoma.

3.4. Mbeya Region

$$Z_1 = 0.00056A - 0.0013B + 0.00044C + 0.0055D \quad 9$$
$$Z_2 = 0.00044A + 0.00204B - 0.000196C - 0.00134D \quad 10$$
$$Z_3 = 0.0036A - 0.00048B - 0.00059C + 0.0028D \quad 11$$

According to Z1, D had the largest influence on distinguishing the samples compared to B, A, and C, capturing 87.96% of the variants. Z2 and Z3 accounted for 10.04 percent and 1.99 percent of the total discrepancy, respectively. From Figure 5, diarrhoea and influenza have a significant impact on B, whereas UTIs with lower total variances were associated with a higher chance of affecting D, C, and A. The study discovered that 92.7% of the disease-specific samples were accurately categorized in their relative covariates with a misclassification rate of 7.3%.

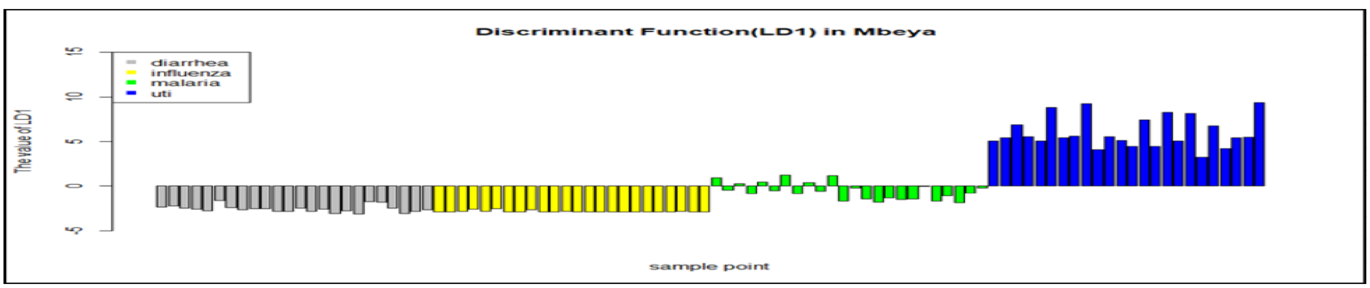

Figure 4: Displaying the results of the first discriminant function for disease incidences in Mbeya.

3.5. Mtwara Region

$$Z_1 = 0.00044A - 0.00093B + 0.00055C + 0.0037D \quad 12$$
$$Z_2 = 0.00021A - 0.00056B + 0.00034C - 0.00134D \quad 13$$
$$Z_3 = -0.0041A + 0.000315B + 0.00079C - 0.0035D \quad 14$$

The study demonstrates that D had the greatest influence on classifying the samples, as shown by Z1. D, then B, C, and A, fared far better than the discriminant function, respectively. Z1 effectively accounts for 94.7%, followed by Z2, and Z3 contributes 3.9% and 1.4% of the total variance. As seen in Figure 5, diarrhoea and influenza have a negative impact on covariate B. Malaria and UTI also have a higher risk for D, C, and A. The study reveals that 8.3% of the samples were misclassified, while 91.7% of the experiment units are accurately categorized into their related covariate.

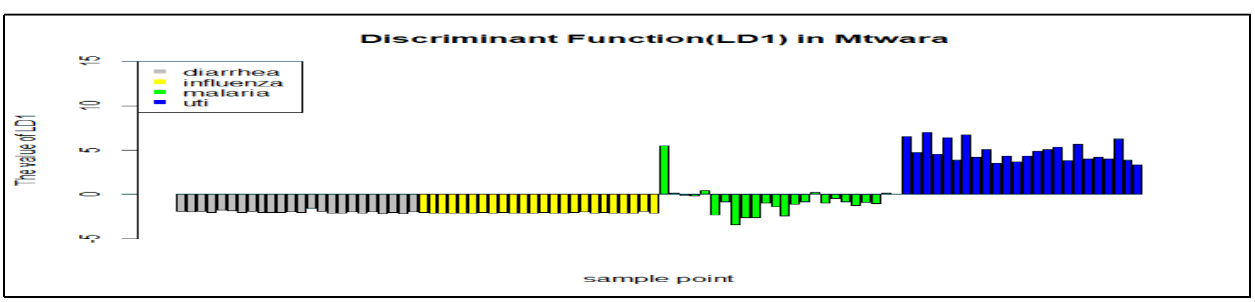

Figure 5: Displaying the first discriminant function for disease incidences in Mtwara.

3.6. Mwanza Region

$$Z_1 = 0.0025A - 0.0016B + 0.0007C + 0.00185D \quad 15$$

$$Z_2 = -0.00081A + 0.00078B - 0.00018C + 0.00032D \quad 16$$

$$Z_3 = 0.361A - 0.0727B - 0.0495C - 0.044D \quad 17$$

For Z1, the risk associated with the covariate is greater for A than for D, B, and C. Z2 and Z3 account for 22% and 2.22% of the overall variance, respectively. 87.57 percent of the variations between the diseases were accurately captured by Z1. The results suggest that variable C has the greatest impact on both diarrhoea and influenza. As illustrated in Figure 6, the results of the first discriminant function for disease incidence demonstrate that variables B and D have been favourably influenced by UTI and malaria. The study reveals that 100% of the samples collected from various diseases are correct to their appropriate variables.

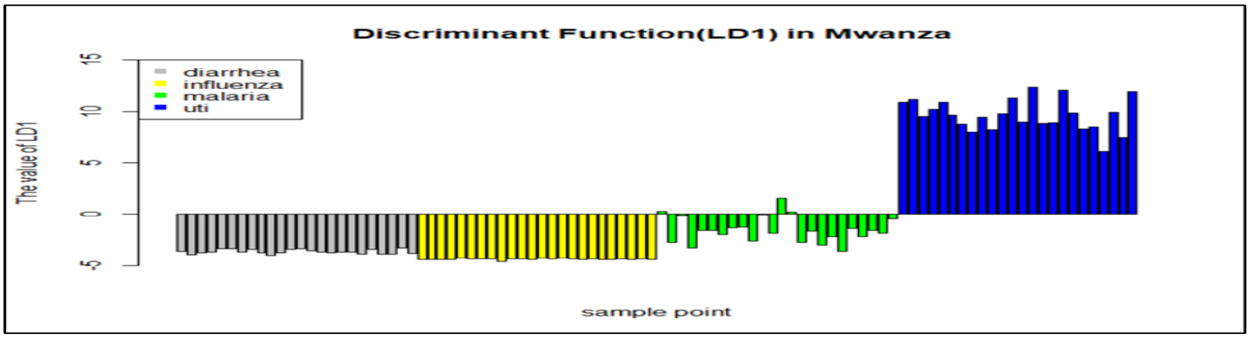

Figure.6: Displaying the first discriminant function for disease incidences in Mwanza.

4. Dendrogram, Cluster Membership, and Silhouette

4.1. Dendrogram and Cluster Membership

The dendrograms, as portrayed in Figures 7 to 12, show the outcomes of both the agglomerative and divisive methods in determining the clusters of the different zones. The study illustrated each cluster's membership and behaviour in chronological sequence. The study sought to compare and contrast how each patient is observed by cluster memberships in each zone, as shown in table 1.

Table: 1. Cluster member in Arusha Region

	1	2	3
diarrhea	0	1	23
influenza	0	24	0
malaria	0	24	0
uti	24	0	0

Table: 2.Cluster member in Dar es salaam Region

	1	2	3
diarrhea	24	0	0
influenza	0	0	24
malaria	24	0	0
uti	0	24	0

Table: 3. Cluster member in Dodoma Region

	1	2	3
diarrhea	2	22	0
influenza	0	24	0
malaria	5	19	0
uti	0	0	24

Table: 4. Cluster member in Mbeya Region

	1	2	3
diarrhea	24	0	0
influenza	24	0	0
malaria	12	0	12
uti	0	21	3

Table: 5. Cluster member in Mtwara Region

	1	2	3
diarrhea	0	24	0
influenza	0	24	0
malaria	10	0	14
uti	24	0	0

Table: 6. Cluster member in Mwanza Region

	1	2	3
diarrhea	0	1	23
influenza	0	24	0
malaria	0	24	0
uti	24	0	0

Table 1: Cluster Membership in its Respective Zone

The study demonstrates that the clusters might vary depending on the types of disease occurrences they experience. In a particular region, the clusters are stated as follows.

4.2. Arusha Region

The study showed three clusters in which UTIs and diarrhoea occur independently and form separate clusters, whereas malaria and influenza occur near to one another to form their own cluster. The dendrogram for the Arusha region is depicted in Figure 7, where the first cluster contained 24 observations of UTIs, the second cluster contained 49 observations each of malaria and influenza and only 1 observation of diarrhoea, and the third cluster contained 23 observations of diarrhoea.

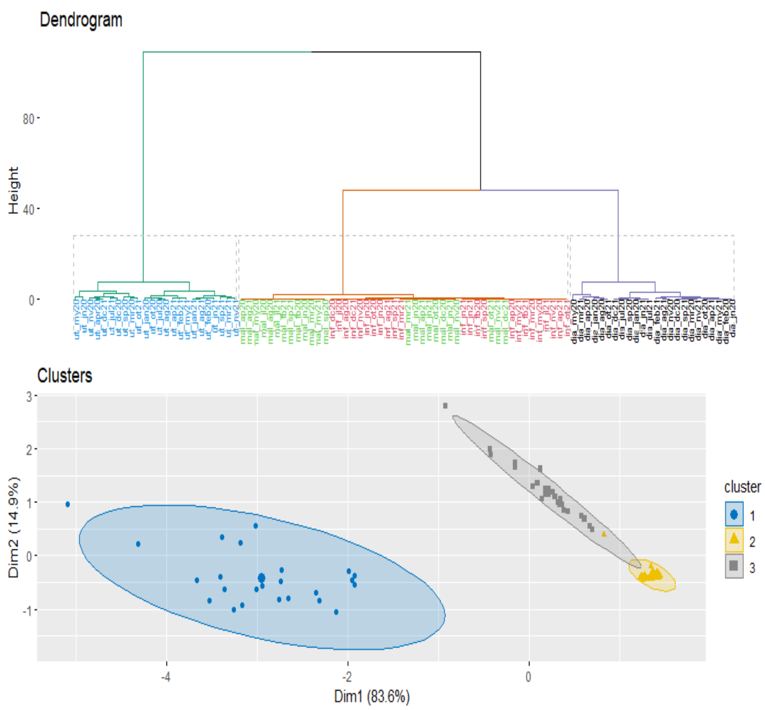

Figure 7: Cluster Dendrogram in Arusha

4.3. Dar es Salaam Region

According to the study's dendrogram in Figure 8, UTIs and influenza are located in the third and second clusters, respectively, whereas malaria and diarrhoea are located in the first cluster, suggesting that they had similar occurrences. The research shows that there were 48 observations in the first cluster but only 24 in each of the other clusters.

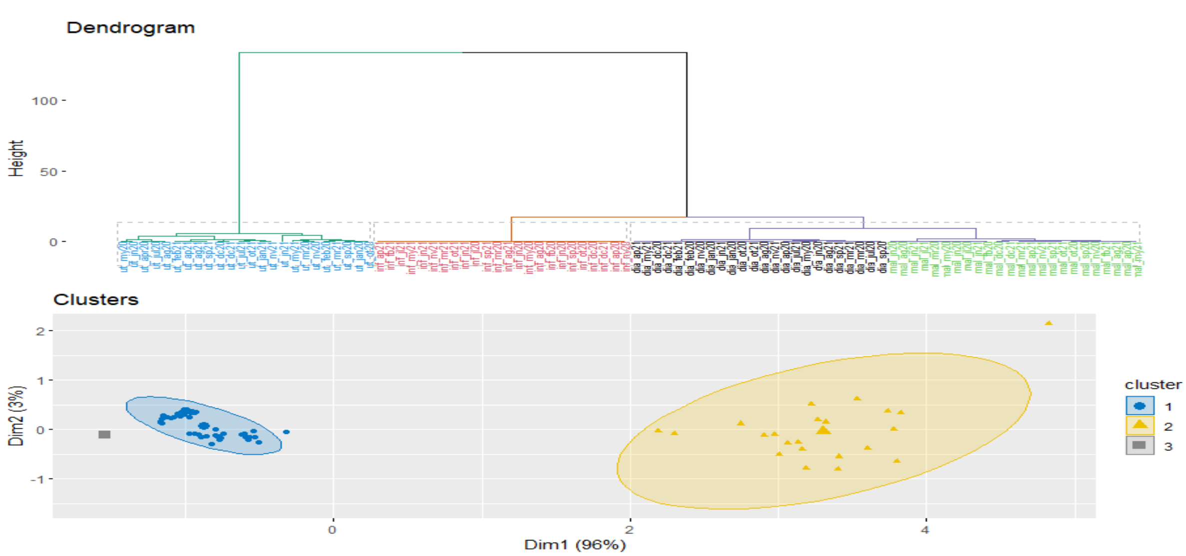

Figure 8: Cluster Dendrogram in Dar es Salaam.

4.4. Dodoma Region

The third cluster, which had 24 observations, is dominated by UTIs, while the second cluster, which had 65 observations, is made up of 24 observations of influenza, 19 observations from malaria, and 22 observations from diarrhoea; and the first cluster, which had seven observations, has two observations each from diarrhoea and five from malaria.

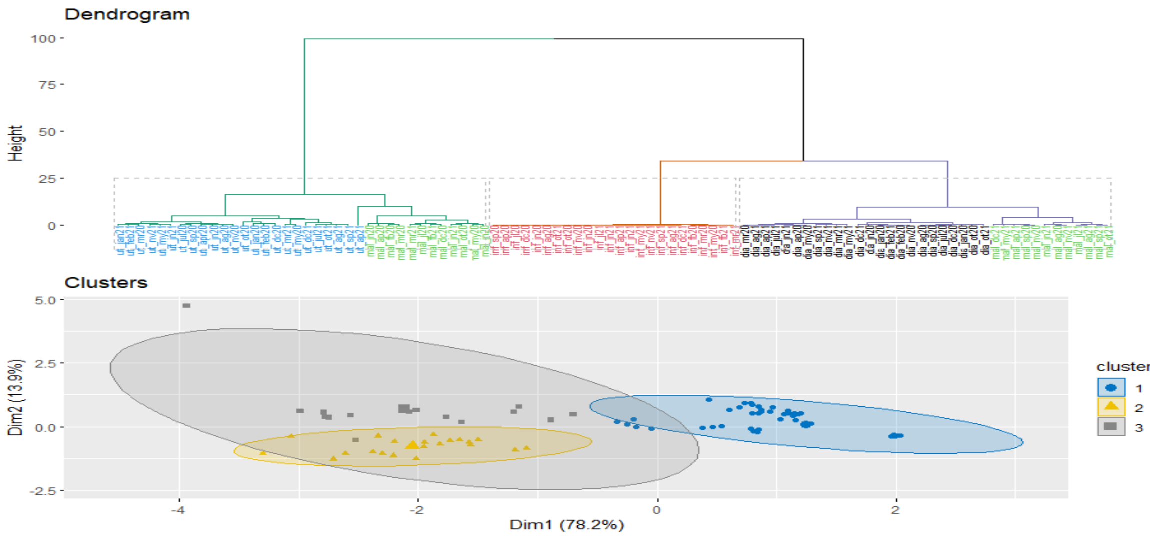

Figure 9: Cluster Dendrogram in Dodoma.

4.5. Mbeya Region

The first cluster had 60 observations, of which 24 are related to diarrhoea, 24 are related to influenza, and 12 are related to malaria. According to Figure 4.23, the second cluster comprised 21 observations from UTIs, while the third cluster contained 15 observations, 3 of which were UTIs and 12 of malaria.

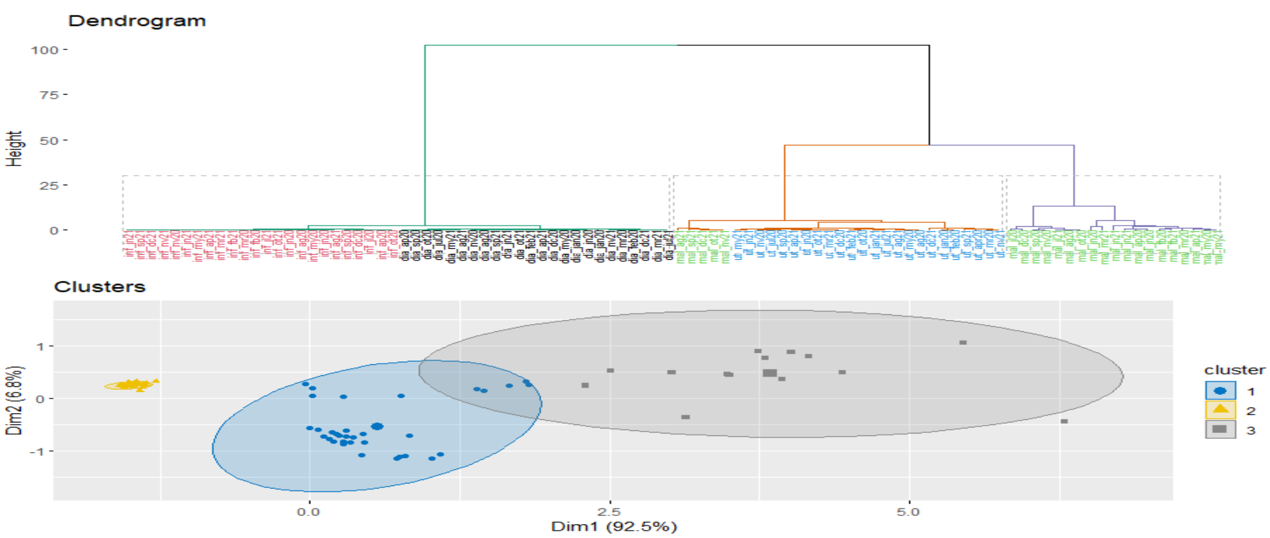

Figure 10: Cluster Dendrogram in Mbeya.

4.6. Mtwara Region

There are 34 observations of malaria and UTIs in the first cluster, of which 10 and 24 observations are related to each. There are also 48 observations in the second cluster, with 24 observations each of influenza and diarrhoea, and there are 14 observations of malaria in the third cluster. This is illustrated in Figure 11.

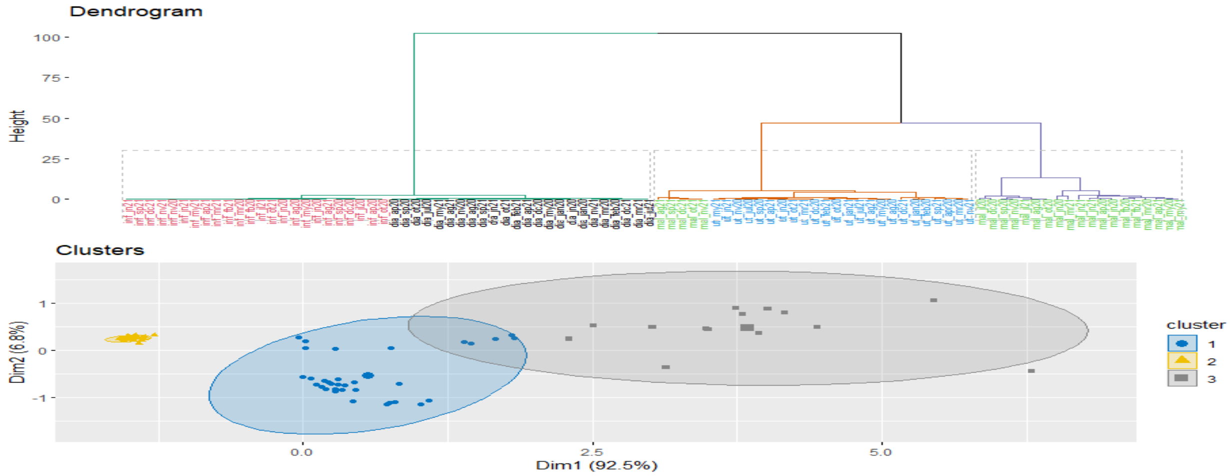

Figure 11: Cluster Dendrogram in Mtwara.

4.7. Mwanza Region

According to the findings, the third cluster included 24 observations related to UTIs; the first cluster had 48 observations, of which 24 observations related to influenza and diarrhoea, respectively; and the second cluster contained 24 observations mostly related to malaria. This is revealed in Figure 12.

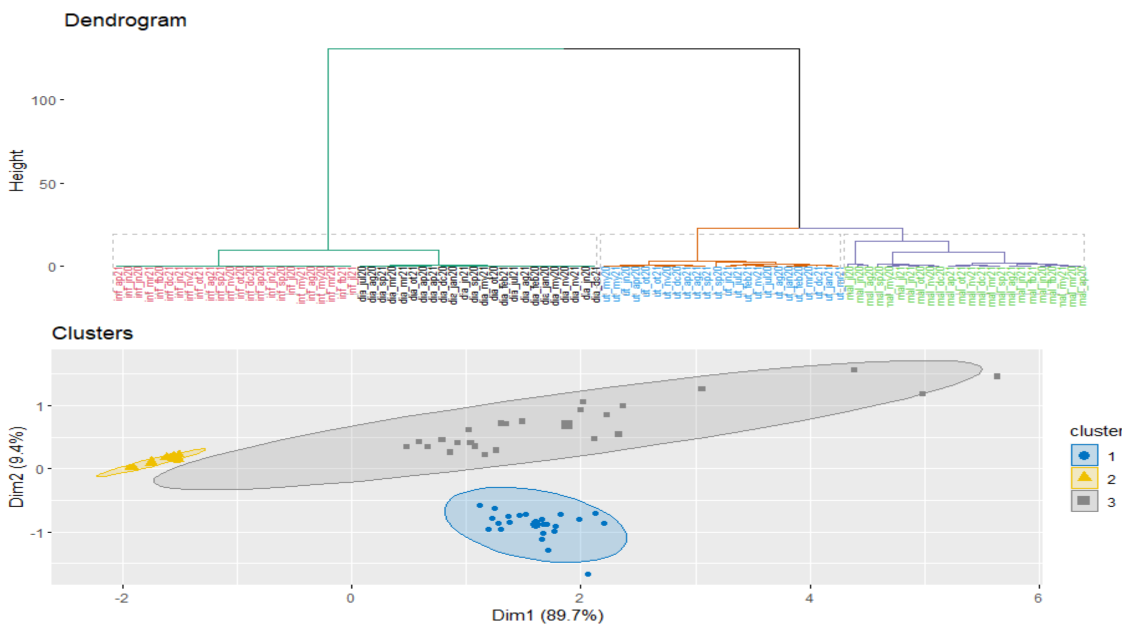

Figure 12: Cluster Dendrogram in Mwanza.

5. Silhouette

The study establishes the average silhouette approach for assessing the level of clustering. It decides the degree to which each observation fits within it's cluster. The study illustrated the average silhouette cluster is a powerful indicator in cluster analysis. The red dot in the illustration represents the average silhouette for each location. Additionally, the findings show each covariate is affected by groups. Therefore, the study proves that in Arusha, the first group is well clustered. In Dar es Salaam, it is also demonstrated that the first group is well clustered. In Dodoma, both of the three groups are well clustered. In addition, clusters two and three in Mbeya are well clustered. As shown in the figures, Mtwara and Mwanza are also well grouped into clusters 1 and 2, respectively.

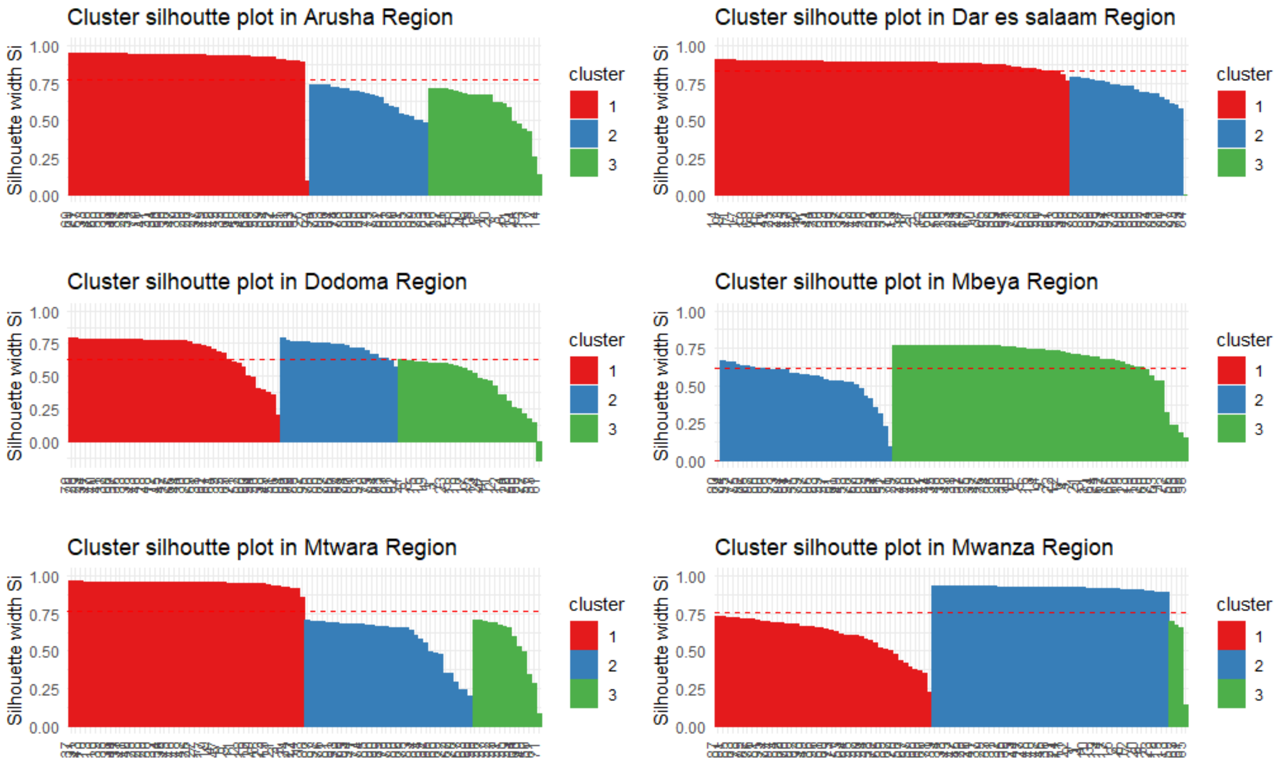

Figure 13: Silhouette of six regions

6. Summary and Discussions

The study is summarized by showing the occurrences of variables in high, moderate, and low categories. The study displayed that due to variation in environmental conditions, the variable category differs from one zone to another.

Regional zone	High	Moderate	low
Arusha	D	A, C	B
Dar es Salaam	B	D, A	C
Dodoma	B	C, D	A
Mbeya	D	B	A, C
Mtwara	D	B	C, A
Mwanza	A	D,B	C

Table 2: Summary of variables by disease risk in respective region

Most of the researchers relied on a univariate approach or, they used linear discriminant analysis as a method based on a single disease for drawing conclusions about the prevalence and incidence rate [20,21]. This study employs multiple diseases using multivariate techniques: discriminant analysis for classifying and recognizing disease groups with high risks. Furthermore, this will help stakeholders in the utilization and allocation of scarce resources in the researched regions. The result given the age group is the following risk arrangement: Those who are at high risk are elders found in the Arusha, Mbeya, and Mtwara regions. Childhood is more risky, in the Dar es Salaam and Dodoma regions, and infants are riskier in the Mwanza region. In addition, at moderate risk, infants are found in the regions of Arusha and Dar es Salaam; elders are found in the Dar es Salaam, Dodoma, and Mwanza regions; and childhood is found in the Mbeya, Mtwara, and Mwanza regions. Also, adolescents found in Arusha and Dodoma are thought to be at moderate risk. Elders were discovered in Dodoma, Mbeya, and Mtwara, as well as adolescents in Dar es Salaam, Mtwara, and Mwanza at low risk and childhood in Arusha is at low-risk.

7. Conclusion

The study concluded the risk recognition at distinct ages is as follows: the elderly are at more risk in the northern, southern highland, and southern zones. Risk is encountered for childhood in the coastal and central areas and infants in the lake zone. In addition, the study advised the government and stakeholders on health system control to make public education plans and assign sufficient resources to prevent calamities. The lakes and northern zones need to have hygiene education and awareness as a strategy for preventing diarrhoea. More resources and strategies for combating malaria should be concentrated in the southern region, and public education about the disease should be intensified. Overall, there is a need for greater resources to be allocated across the Tanzania mainland to combat UTI, which has manifested and predominated throughout zones, as well as extensive public education [22,23].

References

1. Dausey, D. J., Biedrzycki, P. A., Cook, T., Teufel, J., Vendeville, M., & Francis, E. (2016). Planning and response to communicable disease on US domestic air flights.
2. Garcell, H. G., Hernandez, T. F., Baker, E., & Arias, A. V. (1995). Implementation of surveillance systems to determine the burden of communicable diseases in a facility in Qatar. *Information for authors, 1*.
3. The SDGs, sustainable development goals. WHO Library (Issue 2). https://apps.who.int/iris/bitstream/handle/10665/311696/WHO-DAD-2019.1-eng.pdf (2019). World Health Organization. Monitoring Health for the SDGs, Sustainable Development Goals. Global Report, WHO Library, http://apps.who.int/iris (2021)
4. Davgasuren, B., Nyam, S., Altangerel, T., Ishdorj, O., Amarjargal, A., & Choi, J. Y. (2019). Evaluation of the trends in the incidence of infectious diseases using the syndromic surveillance system, early warning and response unit, Mongolia, from 2009 to 2017: a retrospective descriptive multi-year analytical study. *BMC infectious diseases, 19,* 1-9.
5. Islam, M. S., Chaudhuri, I., Al Mobin, M., Islam, M., Mahmud, M. S., KutubUddin, M., ... & Kamrujjaman, M. (2021). The perspective of acquired immunity to combat against infectious diseases: An overview. *Health, 13*(9), 1020-1044.
6. Kibasan, J. I. A., Abufayed, B. F., & Madamba, F. U. (2016). Comparative Study on Awareness of Libyans in the Prevention and Control of Communicable Diseases.
7. Poda, Armel, Jacques Zoungrana, Arsène Héma, Ziemlé Clément Méda, Alexandre Boena, Rainatou Boly, Apoline Sondo et al. "Morbidity and Mortality of Inpatients in the Department of Infectious Diseases of the University Hospital of Bobo-Dioulasso, Burkina Faso." *Advances in Infectious Diseases 9*, no. 3 (2019): 171-182.
8. Wichiern. R. Applied Multivariate Statistical Analysis. Pearson Education (Issue 6). https://www.tandfonline.com/doi/full/10.1080/00224065.1997.11979760 (2007). World Health Organization. World health statistics overview 2019: monitoring health for the SDGs, Sustainable Development Goals. Global Report, WHO Library, http://apps.who.int/iris (2021)
9. Czumbel, I., Quinten, C., Lopalco, P., Semenza, J. C., & ECDC expert panel working group Alberto E. Tozzi Catherine Weil-Oliver Gordon Nichols Hanne Nøkleby Irina Brumboiu Janneke Verheijen Javier Segura del Pozo Mira Kojouharova Emese Szilágyi Tiia Pertel. (2018). Management and control of communicable diseases in schools and other child care settings: systematic review on the incubation period and period of infectiousness. *BMC infectious diseases, 18*, 1-15.
10. Hedlund, C., Blomstedt, Y., & Schumann, B. (2014). Association of climatic factors with infectious diseases in the Arctic and subarctic region–a systematic review. *Global Health Action, 7*(1), 24161.
11. Ndeffo-Mbah, M. L., Vigliotti, V. S., Skrip, L. A., Dolan, K., & Galvani, A. P. (2018). Dynamic models of infectious disease transmission in prisons and the general population. *Epidemiologic reviews, 40*(1), 40-57.
12. Mishra, A., Harichandrakumar, K. T., Satheesh, S., & Nair, N. S. (2021). Multivariate approach in analyzing medical data with correlated multiple outcomes: An exploration using ACCORD trial data. *Clinical Epidemiology and Global Health, 11*, 100785.
13. Kaindoa, E. W., Finda, M., Kiplagat, J., Mkandawile, G., Nyoni, A., Coetzee, M., & Okumu, F. O. (2018). Housing gaps, mosquitoes and public viewpoints: a mixed methods assessment of relationships between house characteristics, malaria vector biting risk and community perspectives in rural Tanzania. *Malaria journal, 17*, 1-16.
14. Mubyazi, G. M., Magnussen, P., Byskov, J., & Bloch, P. (2013). Feasibility and coverage of implementing intermittent preventive treatment of malaria in pregnant women contacting private or public clinics in Tanzania: experience-based viewpoints of health managers in Mkuranga and Mufindi districts. *BMC Health Services Research, 13*, 1-18.
15. Sharman, M., & Bachmann, M. (2019). Prevalence and health effects of communicable and non-communicable disease comorbidity in rural KwaZulu-Natal, South Africa. *Tropical Medicine & International Health, 24*(10), 1198-1207.
16. Khalique, F., & Khan, S. A. (2021). Multiple disease hotspot mining for public health informatics in resource starved settings: study of communicable diseases in Punjab, Pakistan. *IEEE Access, 9*, 89989-89998.
17. Abbas. F.M. Wasin. A.A. Applied Statistics for Environmental Science with R.: Vol. I. Science Direct, Elsevier Publishing, https://shop.elsevier.com/books/applied-statistics-for Environmental-science-with-r/al-karkhi/978-0-12-818622-0 (2020).
18. Huang, L. (2020). Modified hybrid discriminant analysis methods and their applications in machine learning. *Discrete Dynamics in Nature and Society, 2020*(1), 1512391.
19. Boedeker, P., & Kearns, N. T. (2019). Linear discriminant analysis for prediction of group membership: a user-friendly primer. *Advances in Methods and Practices in Psychological Science, 2*(3), 250-263.
20. Gupta, A., Soni, H., Joshi, R., & Laban, R. M. (2022). Discriminant analysis in contrasting dimensions for polycystic ovary syndrome prognostication. *arXiv preprint arXiv:2201.03029*.
21. Rasheed, J. (2022). Analyzing the effect of filtering and feature-extraction techniques in a machine learning model for identification of infectious disease using radiography imaging. *Symmetry, 14*(7), 1398.

22. Sharman, M., & Bachmann, M. (2019). Prevalence and health effects of communicable and non-communicable disease comorbidity in rural KwaZulu-Natal, South Africa. *Tropical Medicine & International Health, 24*(10), 1198-1207.
23. Fredj, M. B., Belguith, A. S., Abroug, H., Dhouib, W., El Mhamdi, S., Salah, A. B., ... & Soltani, M. S. (2017). Hospitalizations for communicable diseases in a developing country: prevalence and trends—Monastir, Tunisia, 2002–2013. *International Journal of Infectious Diseases, 55*, 102-108.