P-ISSN: 2790-6795	Annals of Human and Social Sciences	Jan-Feb 2026,Vol. 7, No.1
O-ISSN:2790-6809	http://doi.org/10.35484/ahss.2026(7-I)25	[293-300]

Annals of Human and Social Sciences
www.ahss.org.pk

RESEARCH PAPER

Spatiotemporal Analysis of HIV Incidence in Sindh, Pakistan: A District-Wise Study (2021–2024) Using Getis-Ord Gi*

[1] Prof. Dr. Rafique Ahmed Chandio and [2] Kamran Ahmed Memon

1. Director General, Sindh Bureau of Statistics, Government of Sindh, Pakistan
2. Statistical Officer, Sindh Bureau of Statistics, Government of Sindh, Pakistan

Corresponding Author memonsirkami@gmail.com

ABSTRACT

This study identifies the geographic clustering of HIV incidence and its changes over time within the districts of Sindh during 2021 - 2024. This study has some limitations, firstly, only public sector diagnosis cases of HIV have been used because of the inaccessibility of private sector. Secondly, the population estimates applied throughout the years may affect the accuracy of the incidence rate. Human Immunodeficiency Virus has become one of the remains a major problem for global public health. The number of new HIV cases has increased significantly over the past 15 years. In Sindh, districts differ in terms of socioeconomic conditions, access to healthcare, and unsafe medical practices. The district level data on HIV diagnosed cases were collected from the Sindh Bureau of Statistics. Spatial autocorrelation methods were applied to derive the findings. Global Moran's I was used to evaluate spatial autocorrelation, while Getis-Ord Gi was utilized to pinpoint areas with high and low incidence rates. The study finds the significant spatial clustering between 2021 and 2023 (Moran's I > 0.29, p < 0.01). By 2024, spatial clustering became less significant, indicating wider spread across districts. Jacobabad, Larkana and Shikarpur districts were continuously classified as hot spots. While Badin, Sujawal, and Mirpur Khas appeared as cold spots. These findings may be helpful for policymakers to understand the dynamics of HIV spread and enable more targeted and effective public health interventions.

Keywords: Spatial Analysis, GIS, Moran's I, Getis-Ord Gi*, Spatial Autocorrelation, Hotspot Analysis, Spatiotemporal Analysis, Sindh

Introduction

Human Immunodeficiency Virus (HIV) remains a dangerous virus that destroys cells of the immune system, making it hard to fight against other diseases. The destruction of the immune system leads to the final stage of this infection, identified as acquired immune deficiency syndrome (AIDS). HIV and AIDS remain a serious threat are a serious threat to the health of citizens, families and communities. The World Health Organization is drawing attention to an important public health problem in Pakistan. In the last 15 years, the number of new HIV infections has increased three times, from 16,000 in 2010 to 48,000 in 2024.The National AIDS Control Program reports that during the first nine months of 2024, Pakistan recorded nearly 9,700 fresh HIV cases, which place it among the countries with the fastest increases in HIV rates in South Asia. This unexpected increase brings attention to specific weaknesses in understanding, precautionary measures, and healthcare facilities, which could potentially impact global health systems. The figure of HIV cases in Sindh varies significantly from district to district, showing that the disease burden is not constant and is influenced by variables such as healthcare accessibility, community knowledge, socioeconomic conditions, and sanitation levels in each district. unsafe medical practices, incorrect blood transfusions, and the frequent use of syringes continue to be major concerns in the rural healthcare system.

The study applied ArcGIS software to analyze the geographical spread of HIV incidence among Sindh districts between 2021 and 2024. The main objectives include the classification of regions with the highest rate of transmission, the determination of significant clusters and the identification of geographical patterns affecting the spread of diseases. Through visualization of these trends, the research aims to provide valuable information to health professionals and administrators to create a targeted and effective intervention. Furthermore, the aim of this study is to spatially identify hot and cold spots of HIV incidence using GIS-based Getis-Ord Gi*statistics.

Study Area

The study area is Sindh, the southern province of Pakistan, which has a population of about 55.7 million in 2023, making it the third-largest province in terms of land area and the second-largest after Punjab. It borders the Arabian Sea to the south, the Indian states of Gujarat and Rajasthan to the east, the provinces of Baluchistan to the west, and Punjab to the north. Administrative divisions of Sindh consist of several districts, each of which has distinguished variations in population density, socioeconomic conditions, climate, and access to healthcare services. For the purpose of this research, a total of 24 districts of the province of Sindh were taken into consideration. The Karachi Division, which is composed of a number of administrative districts, was treated as a single district unit because of population reporting and data aggregation.

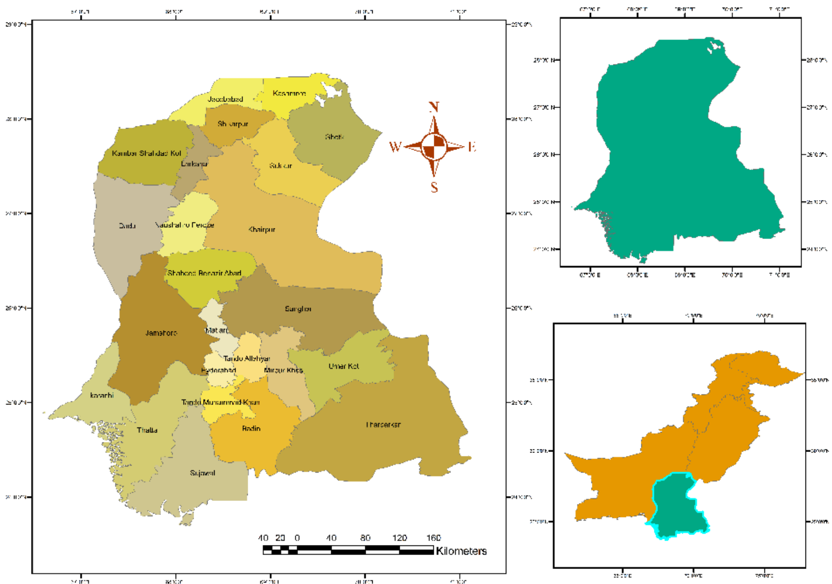

Figure 1. Map of Study Area (Sindh) and its Districts Boundaries

Literature Review

Previous research has reveals that spatial analysis is vital for understanding HIV patterns, as it aids in identifying geographic clustering, measuring spatial variation, and supporting evidence-based planning for effective intervention strategies.

Awoke Seyoum Tegegne et al. (2024) used DHIS2 data (2018-2024) to examine the spatiotemporal distribution of HIV indicators in northwest Ethiopia. Using spatial interpolation and regular Kriging in SPSS analysis, it found that more than 8.5 million participants had HIV positive status and the trend was declining in tests. In many urban areas, key locations were recognized, showing elevated levels of HIV and a rise in the commencement of antiretroviral therapy. The study emphasizes the need of targeted, regionally specific initiatives and shows how successful HIV management strategies may be achieved through the analysis of geographical and temporal data

Sethunya R. Simela and colleagues (2021) examined HIV prevalence in Botswana by analyzing data from the Botswana AIDS Impact Survey (BAIS V). They applied spatial analysis methods, including Local Moran's I, and found notable geographic variations in HIV rates. Areas with high-risk clusters were mainly located in the northeastern parts of the country, whereas urban areas like Gaborone showed lower prevalence. The research emphasized the impact of socio-demographic factors on the spatial distribution of HIV.

Zulu et al. (2014) performed a spatiotemporal investigation of HIV prevalence in Malawi spanning the years 1994 to 2010, with the application of Geographic Information System. Their findings revealed the existence of persistent hotspot areas, notably concentrated in the southern region of the country, whereas the central regions exhibited comparatively lower risk levels. Furthermore, the study recognized critical determinants influencing HIV prevalence, including educational attainment, accessibility to healthcare services, and proximity to transportation infrastructure.

Using GIS-based spatial analysis, this study aims to fill this gap by examining the clustering patterns, and temporal trends of HIV incidence across the districts of Sindh during 2021-2023.

Material and Methods

Data source

In this research, secondary data concerning the annual number of HIV diagnoses by district, sourced from the Sindh Bureau of Statistics for the years 2021 to 2024 was used. All diagnosed cases reported were from public healthcare facilities across each district in Sindh Province. Whereas population data was obtained from the Sindh statistics published by the Bureau of Statistics, Government of the Sindh. The district-wise populations for 2021 and 2022 were projected using the yearly population growth rate that was calculated between the 1998 and 2017 population censuses. For the year 2023, population figures were used directly from the 2023 Population Census, and for 2024, the population estimation for projections was determined using the growth rate between the 2017 Population Census and the 2023 Census. To calculate disease incidence rates for spatial and temporal analysis, these actual and projected population numbers were used as denominators for the respective years. The Incidence rate of each district was calculated for all the selected years by using total number of HIV cases and population for those respective years.

Incidence Calculation

$$\textbf{Incidence Rate} = \frac{\text{Number of reported cases}}{\text{population}} \text{ x } 100{,}000$$

This standardization made it possible to compare data from different districts with different population sizes and from different years

Data Analysis

The annual HIV incidence per 100,000 individuals at each district was determined for the period from 2021 to 2024 and calculated using Microsoft Excel software. The incidence rate is calculated as the number of new cases divided by the total population of each district and multiplied by 100,000. Finally, the data was stored in CSV format and exported to ArcGIS version 10.8.2 for further analysis. ArcGIS version 10.8.2 software was used to analyze Global Moran's I to determine whether the HIV incidence rate is randomly distributed, clustered, or dispersed. In addition, an analysis of hotspots was conducted using the Getis-Ord Gi statistic to identify statistically significant clusters hotspots and cold spots rates of HIV occurrence. Choropleth maps for HIV incidence were created to illustrate spatial trends and changes over time throughout the districts.

Temporal Analysis

The district-level incidence rates of Human Immunodeficiency Virus (HIV) were examined using spatial statistical methods in ArcGIS version 10.8.2. Choropleth maps were created to illustrate the spatial distribution of disease incidence throughout the districts of Sindh. These maps served to highlight differences in the burden of disease, pinpoint geographic clustering patterns, and analyze spatial disparities in disease incidence over the study period.

Spatial Autocorrelation Analysis

Global Moran's I

The Moran I test is the global test most generally used for measuring spatial autocorrelation.

- Null hypothesis (H_0): Spatial distribution is random
- Alternative hypothesis (H_1): Spatial clustering exists

Significance was evaluated using z-scores and p-values ($\alpha = 0.05$).

The neighborhood matrix is used as the foundation for the Moran's I test. which defines the connection between geographical entities. Moran's I values are between -1 and +1. Using Moran's I, z-score, and p-value can also be measured. The positive Z score value shows that neighboring features also have identical data, and the negative Z score indicates dispersion, so these features show outliers.

The formula for Moran's I is expressed as

$$I = \frac{n}{s_0} \cdot \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} wij(x_{i-\bar{x}})(x_{J-\bar{x}})}{\sum_{i=1}^{n} wij(x_{i-x})^2}$$

where n is the number of the samples; x_i is the attribute value of spatial unit i; $\bar{x}$ is the mean value; w_{ij} is spatial weight; S_0 is the sum of all elements of the spatial weight matrix

Hotspot Analysis Using Getis-Ord Gi*

The Getis-Ord Gi measure, a commonly utilized local indicator of spatial association (LISA), was used to evaluate the existence and strength of spatial grouping of HIV rates. This method identifies statistically significant high-value (hot spots) and low-value (cold spots) groupings. Hotspot areas of HIV cases across different districts were identified using the Getis-Ord Gi technique. Different levels of confidence were indicated by the evaluation's Gi* Z-scores and Gi-Bin values, which ranged from −3 to +3. Specifically, Z-score boundaries of ±1.65, ±1.96, and ±2.58 correspond with confidence intervals of 90%, 95%, and 99%.

Mathematical Basis and Definition

The Getis-Ord Gi* statistic evaluates, for each spatial unit i, whether the attribute values in its local neighborhood are significantly higher or lower than would be expected under global spatial randomness. The formula for Getis-Ord G* is given as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j \; - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{n\sum_{i=1}^{n} w^2{}_{ij} - (\sum_{j=1}^{n} w_{i,j})^2}{n-1}}}$$

In the equation, X is the mean of all the values, and S is the standard deviation. The x_j is the value of the attribute at j_{th} location w_{ij} is the value of spatial weight attributed to the i and j location.

This uniform statistic follows a normal distribution under the null hypothesis of spatial randomness.

- A high positive G_i^* indicates a statistically significant hotspot.
- A large negative G_i^* indicates a statistically significant cold spot.

Results and Discussion

Spatiotemporal Trends of HIV Incidence Rates in Sindh, Pakistan (2021–2024)

From 2021 to 2024, the distribution of HIV incidence rate across Sindh's districts is shown in Figure 2. The study divides the districts into very low, low, medium, high, and very high incidence rates using a quantile-based method. In 2021, Larkana and Jacobabad had the highest incidence rates, whereas Shikarpur and Hyderabad also had significant high values, while most other districts were categorized as either very low or low risk. In 2022, the situation deteriorated, with Jacobabad moving into the very high incidence category and several other districts shifting from very low to low or moderate levels. By 2023, more districts were classified as moderate, and the HIV incidence increased already identified in high areas. In 2024, Hyderabad was classified as very high risk, the Mirpurkhas district appeared as a new hotspot, Numerous districts that were formerly categorized as low or moderate now report a rise in cases. This indicates that HIV is progressively spreading and increasing within Sindh over time.

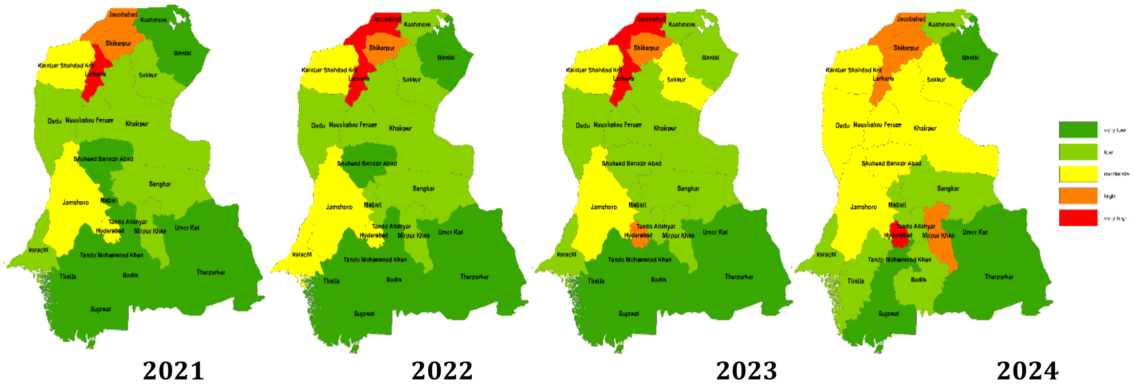

Fig. 2. Spatiotemporal Distribution of HIV Incidence Rates in Sindh

Table 1

Spatial Autocorrelation Analysis of HIV Incidence Using Global Moran's I Index

Year	Moran's I	z-score	p-value	Pattern
2021	0.338982	3.339	0.000841	Significant Clustering
2022	0.368568	3.794466	0.000148	Significant Clustering
2023	0.297951	2.970882	0.002969	Significant Clustering
2024	0.031134	0.622670	0.533501	Not Significant Clustering

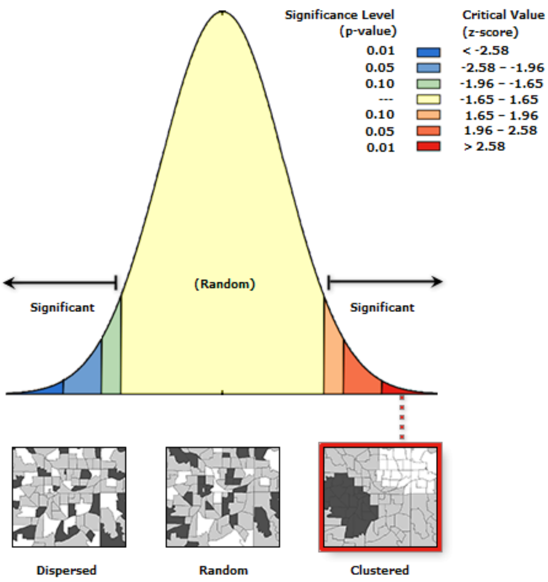

Fig. 3. Interpretation of Getis-Ord Gi* Z-Scores and Statistical Significance

The results from the spatial autocorrelation analysis using global Moran's I are shown in Table 2. This analysis was aimed to analyze whether the rates of HIV incidence were clustered in specific areas, distributed randomly, or spread out across the districts in Sindh. The values for Moran's I for the years 2021 (0.3389), 2022 (0.3686), and 2023 (0.2979) were all positive and statistically significant, with z-scores above 2.9 and p-values less than 0.01. These results suggest a strong clustering of HIV occurrences in these years, showing that regions with high HIV infection rates were located near other areas with similarly high rates. The year 2022 showed the most significant spatial clustering, as Moran's I reached 0.3686 and the z-score was 3.79, indicating a robust spatial relationship among neighboring districts. However, in 2024, the Moran's I value significantly dropped to 0.031, with a z-score of 0.62 and a p-value of 0.53, which revealed that the spatial arrangement of HIV cases in that year lacked any statistical significance. This suggests that HIV cases became more spatially dispersed across districts rather than concentrated in specific clusters. These findings indicate that while HIV incidence exhibited strong spatial clustering during the early years of the study period, the pattern became more spatially dispersed by 2024.

Table 2
Spatial Hotspot Analysis of HIV Incidence in Sindh (2021–2024)

Year	Hotspot Districts (High HIV Prevalence)	Cold Spot Districts (Low HIV Prevalence)	Districts with No Significant Clustering
2021	Jacobabad (99%), Shikarpur (99%), Larkana (99%), Kamber Shahdadkot (99%), Dadu (90%)	Badin (90–95%), Sujawal (90–95%)	Remaining districts (not significant)
2022	Jacobabad (99%), Shikarpur (99%), Larkana (99%), Kamber Shahdadkot (99%), Kashmore (95–99%)	Badin (90–95%), Mirpur Khas (90%)	Remaining districts (not significant)
2023	Jacobabad (99%), Shikarpur (99%), Larkana (99%), Kamber Shahdadkot (99%)	Badin (90%), Mirpur Khas (90%)	Remaining districts (not significant)
2024	Jacobabad (90%), Shikarpur (90%), Larkana (99%), Kamber Shahdadkot (90%)	Sujawal (90%)	Remaining districts (not significant)

Table 02 reveals the spatial distribution of HIV hotspots and cold spots in Sindh. Between 2021 and 2024, clusters of HIV hotspots with confidence levels as high as 99% were consistently found in the districts of Jacobabad, Shikarpur, Larkana, and Kamber Shahdadkot. This suggests that there is a strong pattern of high HIV prevalence in these districts. In the years 2021 and 2022, Dadu and Kashmore are detected as statistically significant areas with a concentration of hotspots with 90% confidence. On the other hand, districts such as Badin, Sujawal, and Mirpur Khas showed low value clusters with a 90% confidence level. Meanwhile, the other districts did not show any clear pattern, which means HIV cases there were spread randomly and not grouped in any specific district.

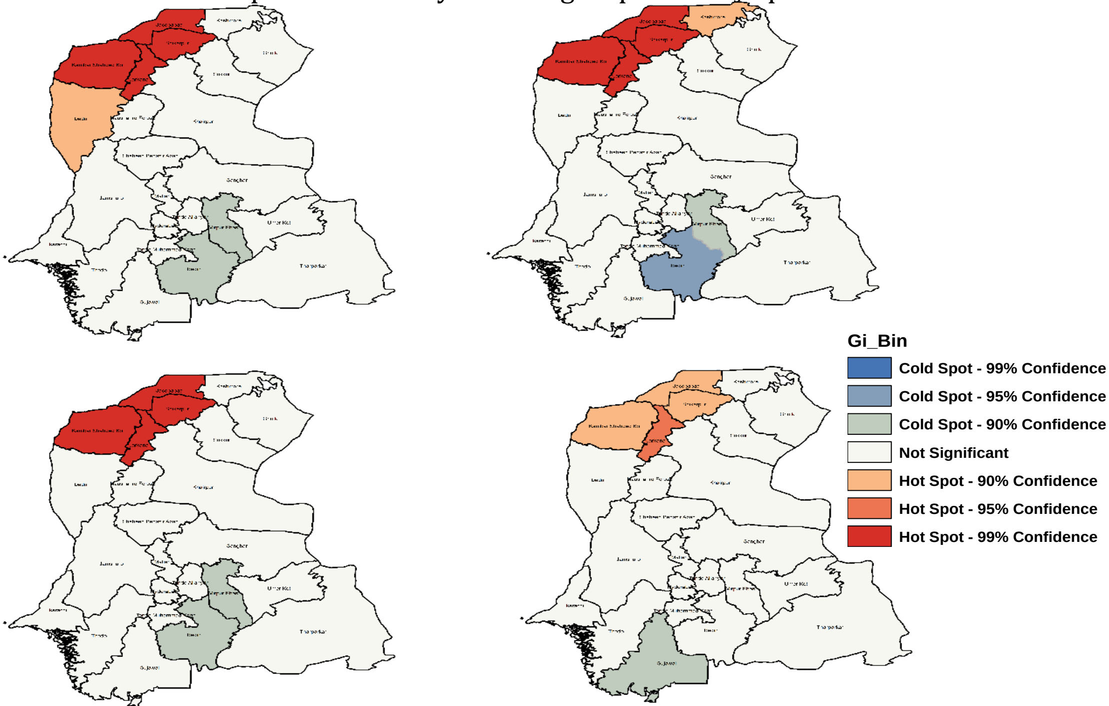

Fig. 4. Spatial clustering of HIV incidence in Sindh Province 2021-2023, based on the Getis–Ord Gi* statistic.

Conclusion

This study analyzes the spatial and temporal clusters of HIV incidence in the Sindh province of Pakistan from 2021 to 2024 using spatial statistics based on the Geographic Information System (GIS). With the application of global Moran's I and Getis-Ord Gi*, the study successfully identified global spatial dependence and local clustering patterns of HIV incidence. This shows that HIV cases were strongly clustered in some regions in2021,2022 and 2023. This means that the high HIV rates are located near other similar high-risk districts, which are clear hotspots. In particular, Jacobabad, Larkana, Shikarpur and Kamber Shahdadkot have been identified with a high level of confidence as areas of high risk. While districts such as Badin, Sujawal, and Mirpur Khas were identified as cold spots, showing relatively lower HIV incidence and maybe better control measures or differing socio-demographic conditions. Hotspots showed a decrease in strength with a 90% certainty for the majority of districts, and only Larkana still showed high significance (99%). Sujawal was identified as a cold spot. This finding matches with the Moran's I result, which was not significant, indicating a more random distribution in 2024. From a policy viewpoint, these findings highlight the persistent need for focused and evidence-based measures in the hotspot areas of Sindh. The measures should include increasing access to HIV testing and screening facilities, launching thorough public education and awareness campaigns, encouraging safe medical and injection practices, and enhancing health infrastructure in at-risk communities.

Recommendations

- Increasing access to HIV testing and screening facilities in hotspot areas
- Launching public education and awareness campaigns
- Encouraging safe medical and injection practices
- Enhancing health infrastructure in at-risk communities

References

Anselin, L. (1995). Local indicators of spatial association—LISA*. Geographical Analysis, 27*(2), 93–115

Boyda, D. C., Holzman, S. B., Berman, A., Grabowski, M. K., & Chang, L. W. (2019). Geographic information systems, spatial analysis, and HIV in Africa: A scoping review. *PLOS ONE, 14*(5), e0216388.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*(3), 189–206.

National AIDS Control Programme. (2024). *Pakistan HIV surveillance report.* National AIDS Control Programme

Raza, H. A., Raja, M. H. R., Khakwani, M. M., & Jamil, B. (2024). Pakistan's HIV high-risk populations: Critical appraisal of failure to curtail spread beyond key populations. *International Journal of Infectious Diseases and Regional Health*, Article 100364.

Simela, S. R., Kelepile, M., & Sebobi, T. I. (2025). Spatial analysis and associated risk factors of HIV prevalence in Botswana: Insights from the 2021 Botswana AIDS Impact Survey (BAIS V). *BMC Infectious Diseases, 25*, Article 69.

Sindh Bureau of Statistics. (2023). *Sindh statistics publication*. Government of Sindh.

UNAIDS. (2023). *UNAIDS data book* 2023. UNAIDS

World Health Organization. (2025). *HIV/AIDS fact sheet 2025*. World Health Organization

Zulu, L. C., Kalipeni, E., & Johannes, E. (2014). Analyzing spatial clustering and the spatiotemporal nature and trends of HIV/AIDS prevalence using GIS: The case of Malawi, 1994–2010. *BMC Infectious Diseases, 14*, Article 285.