

Identification of spatial clusters and spatial outliers of HIV/AIDS using local Moran's I statistic

Thi-Hien Cao * and Thi-Hong Dao

Faculty of Nursing, East Asia University of Technology, Hanoi, Vietnam.

World Journal of Biology Pharmacy and Health Sciences, 2024, 17(01), 101-108

Publication history: Received on 30 November 2023; revised on 16 January 2024; accepted on 19 January 2024

Article DOI: <https://doi.org/10.30574/wjbphs.2024.17.1.0012>

Abstract

Background: Human immunodeficiency virus (HIV) remains a major global public health issue, having claimed 40.4 million lives so far with ongoing transmission in all countries globally; with some countries reporting increasing trends in new infections when previously on the decline. This study aims to use the local Moran's I statistic to investigate the spatial clusters and spatial outliers of HIV/AIDS cases 63 provinces/cities in Vietnam in 2017.

Methods: The local Moran's I statistic is first employed to measure spatial auto-correlation between the number of HIV/AIDS cases in each province/city, and then identify the spatial clusters and spatial outliers of HIV/AIDS cases. Spatial distribution of HIV/AIDS clusters and outliers will be mapped with the help of a GIS. Finally, the main findings will be discussed and summarised.

Results: High numbers of HIV/AIDS cases were mainly concentrated in the provinces of the north central region, Da Nang and provinces in the south of Vietnam. Low numbers of HIV/AIDS cases were detected in the northeastern provinces, central and southeastern provinces of Vietnam. Specifically, one high-high cluster and six low-low spatial clusters, and four low-high and high-low spatial outliers of HIV/AIDS cases were successfully detected. Whereas, the only high-high spatial cluster was discovered in Binh Duong province with 3598 HIV/AIDS cases.

Conclusions: Local Moran's I statistic can help to effectively identify spatial clusters and spatial outlier of HIV/AIDS. Findings in this study provide an insight into how to use spatial statistics to study the spread of HIV/AIDS.

Keywords: Local Moran's I statistic; Spatial outliers; Spatial clusters; Spatial distribution, HIV/AIDS; Vietnam

1. Introduction

HIV, the causative pathogen of AIDS, was first reported in 1981 (1) and has been recognized as one of the most lethal infectious diseases globally for decades, with a serious impact on public health (2). Human immunodeficiency virus (HIV) is an infection that attacks the body's immune system. Acquired immunodeficiency syndrome (AIDS) is the most advanced stage of the disease. HIV targets the body's white blood cells, weakening the immune system. This makes it easier to get sick with diseases like tuberculosis, infections and some cancers. HIV is spread from the body fluids of an infected person, including blood, breast milk, semen and vaginal fluids. It is not spread by kisses, hugs or sharing food. It can also spread from a mother to her baby. In 2016, HIV/AIDS directly cost 57.6 million disability-adjusted life years (DALYs) and 44.8 million DALYs in related syndromes caused by HIV/AIDS (2,3). In 2020, globally, 36 million adults over the age of 15 were living with HIV (4). Out of these, 84% knew their status, 73% were accessing treatment and 66% were virally suppressed (4). South Africa contributes approximately 22% of the global HIV burden (4,5), and KwaZulu-Natal province is the epicentre (6,7), where the UNAIDS targets has not been met (8,9). Whilst South Africa

* Corresponding author: Thi-Hien Cao

has substantially scaled-up ART provision, having the largest HIV treatment programme globally, has resulted in reducing number of HIV related death (2). However, country level HIV prevalence of 14.0%, with an estimated 231,000 new infections remains persistently high (10), and almost a fourth of women in their reproductive ages (11,12) were HIV positive at the end of 2020 (2). However, data from a recent report showed that there were an estimated 39.0 million people living with HIV at the end of 2022, two thirds of whom (25.6 million) are in the WHO African region.

Spatial analysis is widely used in HIV/AIDS research to identify high-risk and spatiotemporal clusters, assess the geographical distribution of infections, and explore the spatial relationship between HIV/AIDS and social factors (13,14). Spatial analysis is a powerful tool to understand the long-term trends and spatial clustering of HIV/AIDS cases. Spatial analysis can provide scientific perspectives for public health professionals and policymakers to design targeted countermeasures (2). In addition, geospatial analytical methods, including geographic information systems (GIS), are an essential tool for understanding the nature and causes of spatial variation in HIV prevalence (15). Yet even with the significant increase in the use of such geospatial tools in understanding public health problems in planning and implementing interventions and assessing their outcomes, geographically explicit studies of HIV/AIDS in sub-Saharan Africa are still very limited (16,17). Reasons for limited GIS use include scarcity of reliable spatially coded data (15). Nykiforuk and Flaman identify four categories of GIS use from a review of 621 cases published between 1990 and 2007, namely: disease surveillance, risk analysis, access to health services and planning, and profiling community-health service utilization (18).

In Vietnam, the first case of HIV infection was reported in December 1990 in Ho Chi Minh City. By 1992, only 11 cases had been reported, but in 1993 there was a sharp increase and twelve years later, in December 2005, the cumulative number of reported HIV-infected people from the 64 provinces had grown to 104,111. Of these, 13,731 were new infections, 17,289 were AIDS patients and 10,071 had died (19). Estimates put the actual number of infections much higher. The HIV epidemic in Vietnam is predominantly drug-related; injecting drugs users have accounted for most (53%) of the recorded infections, although this data from surveillance may be incomplete. HIV prevalence in that group is still increasing, in 2005 at approximately 30%. The epidemic affects mainly young males; 64% of reported cases are men under 29 years of age (20). Understanding the spatial distribution and spatial clustering of HIV/AIDS plays a very important role in the control of HIV/AIDS pandemic. Therefore, this study was carried out to use the local Moran's I statistic to investigate the spatial clusters and spatial outliers of HIV/AIDS cases 63 provinces/cities in Vietnam in 2017. The local Moran's I statistic is first employed to measure spatial auto-correlation between the number of HIV/AIDS cases in each province/city, and then identify the spatial clusters and spatial outliers of HIV/AIDS cases. Spatial distribution of HIV/AIDS clusters and outliers will be mapped with the help of a GIS. Finally, the main findings will be discussed and summarised.

2. Material and method

2.1. Material

HIV prevalence data in Vietnam is based primarily on HIV/AIDS case reporting and on the HIV sentinel surveillance conducted annually in 40 of Vietnam's 64 provinces. The government now reports HIV cases in all provinces, 93 percent of all districts, and 49 percent of all communes, although many high prevalence provinces report cases in 100 percent of communes. Even though Vietnam has implemented HIV/AIDS case reporting, the general lack of HIV testing thus far suggests that the actual number of people living with HIV/AIDS is much higher. The first HIV case in Vietnam was detected in 1990 (19). The estimated number of people living with HIV then rose drastically from 3,000 in 1992 to 220,000 in 2007, and is projected to be 280,000 in 2012. Among these, 5,670 are children. According to the IMF, this trend is placing Vietnam at the threshold of moving the disease from the high-risk groups of drug users and sex workers to the general population (21). Among those who inject drugs, 19% are infected by HIV (up to 30% in some provinces). In this study, datasets of HIV/AIDS cases in 63 provinces/cities in Vietnam in 2017 collected from website of the Vietnam Ministry of Health (VMH) were used to identify spatial outliers and spatial clusters of HIV/AIDS using the local Moran's I statistic.

2.2. Method

Spatial autocorrelation is used to describe the extent to which a variable is correlated with itself through space. This concept is closely related to Tobler's First Law of Geography, which states that "everything is related to everything else, but near things are more related than distant things" (22). Spatial autocorrelation is a fundamental concept in spatial analysis (25). It is the correlation among values of a single variable strictly attributable to their relatively close locational positions on a two-dimensional surface, introducing a deviation from the independent observation assumption of classical statistics (26). Positive spatial autocorrelation means that geographically nearby values of a variable tend to

be similar on a map: high values tend to be located near high values, medium values near medium values, and low values near low values. Positive spatial autocorrelation occurs when observations with similar values are closer together (i.e., clustered). Negative spatial autocorrelation occurs when observations with dissimilar values are closer together (i.e., dispersed) as shown in Figure 1.

Spatial autocorrelation may be indexed, quantified by including an autoregressive parameter in a regression model, or filtered from variables. Spatial autocorrelation can be quantified with indices. Spatial autocorrelation can be assessed using indices that summarize the degree to which similar observations tend to occur near each other over the study area. Two common indices that are used to assess spatial autocorrelation in areal data are Moran’s I statistic (23) and Geary’s C statistic (24).

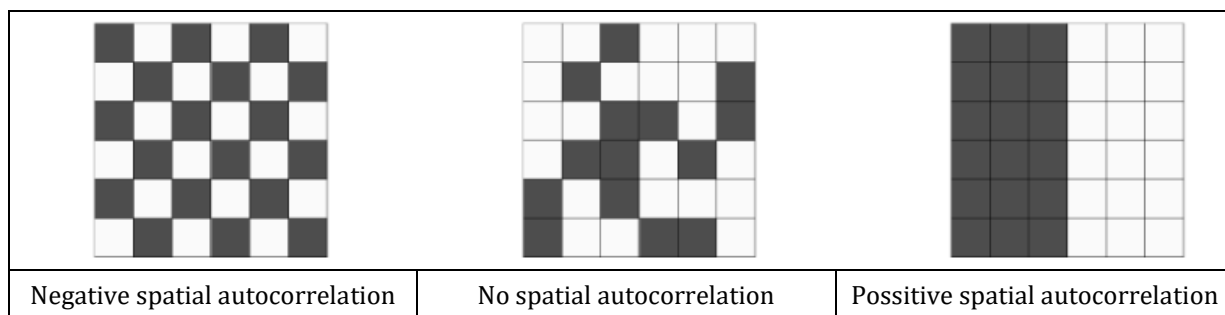


Figure 1 Configurations of areas showing different types of spatial autocorrelation.

This study used global Moran's I statistic to identify the spatial clustering of the HIV/AIDS cases at global scale (27,28). The definition of the global Moran's I statistic is expressed in equation (1):

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^n \sum_{j=1}^n W_{ij} \sum_{i=1}^n (x_i - \bar{x})^2} \dots\dots\dots(1)$$

where x_i and x_j are the number of HIV/AIDS confirmed cases for province/city i and province/city j ; \bar{x} is the mean of HIV/AIDS cases and be given by $\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$; n is the total number of provinces/cities in the whole study area; and W_{ij} is a $(n \times n)$ spatial weight matrix (29).

The range of values of global Moran’s I coefficient is in the interval $[-1, +1]$ (29). Positive values of Moran's I result from the data's positive spatial autocorrelation., whereas Moran's I values are negative when there is a negative spatial autocorrelation (30). The absence of spatial autocorrelation or random HIV/AIDS distribution is shown by values of the global Moran's I coefficients that are near to zero.

The global Moran’s I reflects the presence or lack of spatial autocorrelation as a whole. Previous studies have shown that general spatial autocorrelation can hide the change of local state, so local spatial autocorrelation was used to detect specific local clusters (28). Moreover, even if there is no global autocorrelation or no clustering, we can still find clusters at a local level using local spatial autocorrelation analysis. Therefore, the regional Moran's I statistic was used to quantify the spatial clustering of low and high HIV/AIDS levels in each province/city (29). The local Moran’s I statistic (I_i) for HIV/AIDS at province/city i is given by the following equation (23):

where x_i , x_j , \bar{x} , and W_{ij} are defined in equation (1); N is the total number of neighborhood provinces/cities (29); J_i denotes the neighborhood set of HIV/AIDS confirmed cases at province/city i ; $j \neq i$ implies that the sum of all $(x_j - \bar{x})$ of nearby neighbourhood province/city i but not including x_j ; and σ^2 is the variance of x , given in equation (3). W_{ij} defines neighbor connectivity and can be constructed using first order and second of contiguity.

$$\sigma^2 = \frac{1}{N} \sum_{j=1}^N (x_j - \bar{x}) \dots\dots\dots(3)$$

A positive value for I indicates that a feature has neighboring features with similarly high or low attribute values; this feature is part of a cluster. A negative value for I indicates that a feature has neighboring features with dissimilar values; this feature is an outlier. In either instance, the p-value for the feature must be small enough for the cluster or outlier to be considered statistically significant. The level of spatial clustering of the HIV/AIDS at each province/city is indicated by local Moran's I statistic. Similar to the global Moran’s I statistic, the local Moran’s I value at province/city i (I_i) also ranges between -1 and $+1$. There is no spatial autocorrelation of HIV/AIDS cases if the local Moran's I coefficient at

province/city i equals zero ($I_i = 0$). If $I_i > 0$ then there will be a positive spatial autocorrelation of HIV/AIDS cases (29). If $I_i < 0$ then there will be a negative spatial autocorrelation of HIV/AIDS cases. A high positive I_i shows the province/city i has a similarly high or low number of HIV/AIDS cases as its neighbors and called the “spatial cluster” (30). In this case, when there is a positive local spatial autocorrelation, the local Moran's I statistic indicates two types of spatial clusters for HIV/AIDS cases, including: high-high spatial clusters and low-low spatial clusters. Low-high and high-low clusters are also two forms of spatial outliers that are identified using the local Moran's I statistic when there is a negative local spatial autocorrelation. The cluster/outlier type field distinguishes between a statistically significant cluster of high values (high-high), cluster of low values (low-low), outlier in which a high value is surrounded primarily by low values (high-low), and outlier in which a low value is surrounded primarily by high values (low-high). In this work, with the help of the spatial statistics software, GeoDA, developed by Luc Anselin (31), a randomization test was used to test the significance of spatial autocorrelation statistics. Spatial autocorrelation statistics were generated and tested at the significance of 0.05 using 999 permutations.

3. Results and discussion

3.1. Spatial distribution of HIV/AIDS cases

The data in the boxplot and map in Figure 2 shows the distribution of the number of HIV/AIDS cases in Vietnam in 2017. Data from Figure 2 demonstrate two provinces having extremely high numbers of HIV/AIDS cases were Ho Chi Minh and Hanoi with the corresponding number of cases of 47,303 and 20,724 cases, respectively. Then followed by Son La (8,164 cases), Hai Phong (7,331), Thai Nguyen (6,265), An Giang (6,121 cases) and Dong Nai (6,067 cases). The province having the low numbers of HIV/AIDS cases in 2017 included Quang Tri (204 cases), followed by Quang Binh (206 cases), Phu Yen (245 cases), Ninh Thuan (276 cases), and Kon Tum (330 cases) and Binh Dinh (362 cases), respectively.

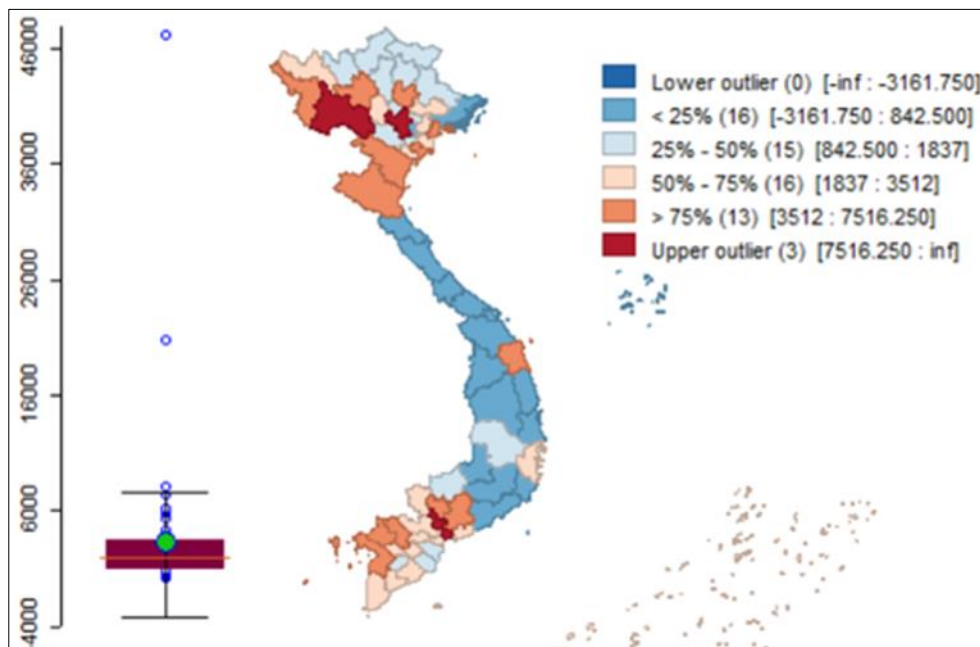


Figure 2 Boxplot and map of HIV/AIDS cases in Vietnam

Data in the boxplot in Figure 2-left shows the distribution of the number of HIV/AIDS cases in Vietnam in 2017. The statistical parameters HIV/AIDS cases in 2017 are as follows: the lowest and highest number of HIV/AIDS cases for each province/city were 204 and 47,303 cases, respectively. The values of mean and median were 3320 and 1837, respectively. It can be seen that the value of the mean was larger than that of the median indicating that the variable of HIV/AIDS tends to skew to the right side of the chart. The statistical values of the first quantile (Q_1) and the third quantile (Q_3) were 842.5 and 3512, respectively. The dispersion values of the interquartile range (IQR) and standard deviation (SD) were 2669.5 and 6384.6, respectively.

The spatial distribution of the number of HIV/AIDS cases in 2017 was shown in the map in Figure 2-right. Data from the map Figure 2-right demonstrates six different ranges of the number of HIV/AIDS cases as follows: (i) smaller than 25 percent in the range of [0, 842], (ii) between 25% and 50% in the range of [842, 1837], (iii) between 50% and 75% in

the range of [1837, 3512], and (iv) upper 75% in the range of [3512, 47303], respectively. The spatial distribution of HIV/AIDS infections in Figure 2-right shows that high numbers of HIV/AIDS cases were mainly concentrated in the provinces of the north central region, Da Nang and some provinces in the south of Vietnam, whereas, low numbers of HIV/AIDS cases were detected in the northeastern provinces, some central and southeastern provinces of Vietnam.

3.2. Analysis of LISA distribution

Data from Figure 3 shows the distribution of LISA determined from HIV/AIDS cases in Vietnam in 2017. Data from the boxplot in Figure 3-left shows the minimum and maximum values of the local Moran's I statistic were -0.54 and 0.40 respectively. The values of the first (1-st quartile - Q1) and third (3-rd quartile - Q3) interquartile ranges were -0.02 and 0.08, respectively. Meanwhile, the Moran's I statistic has an average and median value of 0.11 and 0.03, respectively. The dispersion measurement values of the Moran's I statistic with the interquartile range and standard deviation were 0.10 and 0.14, respectively. It can be seen that the median value was larger than the mean value, therefore, the data on the local Moran's I statistic is skewed toward the top of the boxplot.

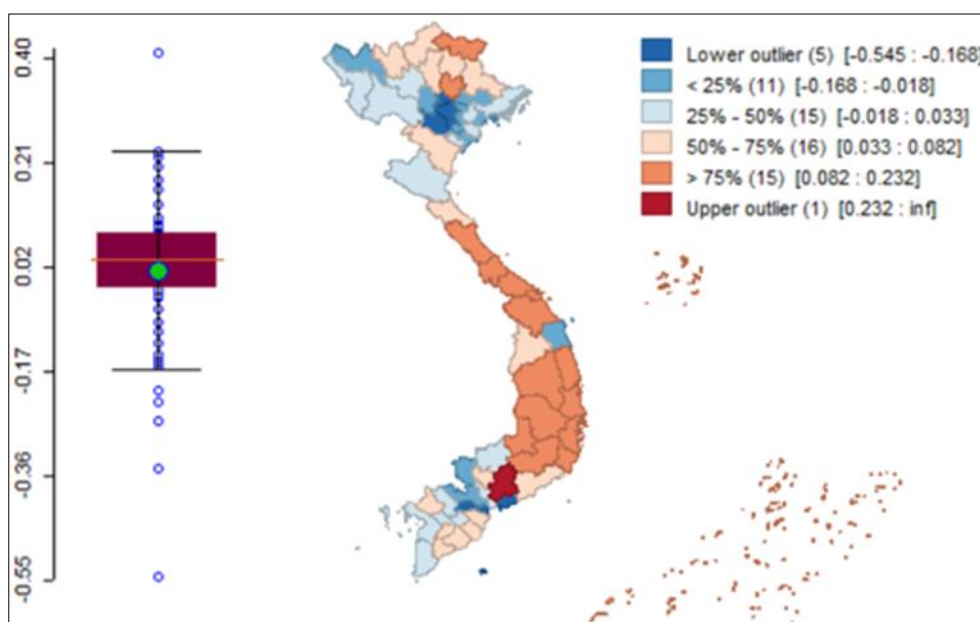


Figure 3 Map of LISA of HIV/AIDS cases in Vietnam

The map in Figure 3-right shows the spatial distribution of the local Moran's I statistic obtained from the variable of HIV/AIDS cases throughout the territory of Vietnam. It can be seen from the map shown in Figure 3 that high values of the Moran's I statistic were mainly detected in the central provinces, some south-central provinces, and the northern provinces of the north of Vietnam. Meanwhile, the local Moran's I statistic was mainly distributed in the northeastern and northwestern provinces, and some provinces in the southwestern region of Vietnam. It can be seen that when combining with the map of HIV/AIDS cases in Figure 3-right, high values of the Moran's I statistic were found in provinces having a large number of HIV/AIDS cases, whereas, provinces having low number of HIV/AIDS cases had a relatively low values of the local Moran's I statistic.

3.3. Analysis of spatial clustering of HIV/AIDS

The results of identifying spatial clusters and outliers of the number of HIV/AIDS cases in Vietnam in 2017 are shown in Figure 4-left. Data from the LISA cluster map in Figure 4-left shows that, in total, there was one high-high cluster and 6 low-low spatial clusters, and 04 low-high and high-low spatial outliers of HIV/AIDS cases were successfully detected throughout Vietnam. The only high-high spatial cluster was discovered in Binh Duong province with 3598 HIV/AIDS cases. Six low-low spatial clusters were found in six provinces having a low number of HIV/AIDS cases such as Quang Binh (206 cases), Quang Tri (204 cases), Thua Thien Hue (362 cases), Dak Nong (405 cases), Lam Dong (804 cases), Khanh Hoa (1853 cases). Low-high spatial clusters were detected in four provinces including Tay Ninh (3109 cases), Long An (2799 cases), Tien Giang (2037 cases) and Ba Ria Vung Tau (2360 cases). The only high-low spatial outliers was identified in Quang Ngai (5,430 HIV/AIDS cases). The map in Figure 4-left also shows that there were no high-high spatial clusters which were detected in some provinces having a high number of cases such as Ho Chi Minh (47,303 cases), Hanoi (20,724 cases), Son La (8,164 cases), Hai Phong (7,331), Thai Nguyen (6,265), An Giang (6,121 cases) and

Dong Nai (6,067 cases). Similarly, low-low spatial clusters and spatial outliers (low-high and high-low) were not found in provinces/cities having a low number of HIV/AIDS infections such as Phu Yen provinces (245 cases), Kon Tum (330 cases) and Binh Dinh (362 cases). In addition, the LISA cluster map in Figure 4-left also shows that the existence of spatial clusters and spatial outliers in 51 provinces/cities was not detected because it does not reach statistical significance at the 0.05 level.

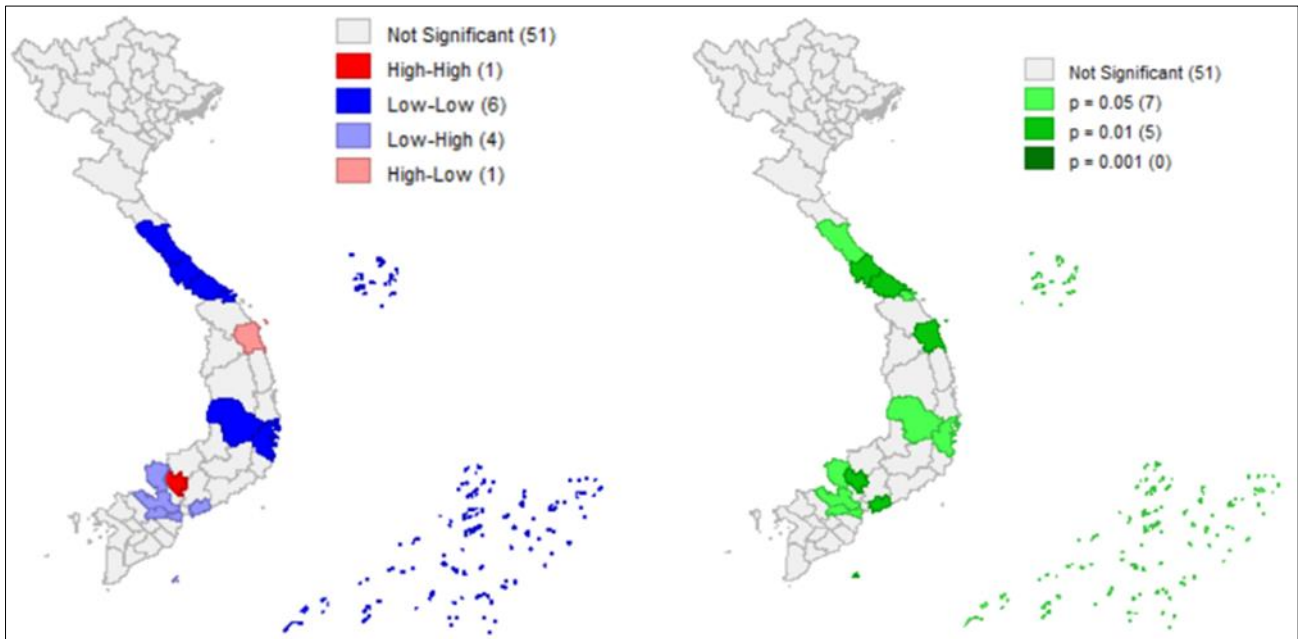


Figure 4 LISA cluster (left) and LISA significant (right) maps of HIV/AIDS cases in Vietnam

Data from Figure 4-right illustrates the spatial distribution of the level of statistical significance (p-value) achieved by the local Moran's I statistic for HIV/AIDS cases collected in each province/city in Vietnam in 2017. Statistical significance levels are expressed on four scales including statistically insignificant (p-value > 0.05) and statistical significance at the levels of 0.05, 0.01 and 0.001. Data from Figure 4-right demonstrate that the local Moran's I statistic in 7 provinces/cities were statistically significant at the 0.05 level, which were Quang Binh (206 cases), Da Nang (778 cases), Dak Lak (1755 cases), Khanh Hoa (1853 cases), Tay Ninh (3109 cases), Long An (2799 cases), and Tien Giang (2037 cases). The local Moran's I statistic found in five provinces were statistically significant at the 0.01 level including Quang Tri (204 cases), Thua Thien Hue (362 cases), Quang Ngai (5430 cases), Ba Ria - Vung Tau (2360 cases) and Binh Duong (3598 cases). In addition, local Moran's I statistic in 51 provinces/cities was statistically insignificant at the 0.05 level (p-value > 0.05).

4. Conclusion

This study was carried out to use the local Moran's I statistic to identify the spatial clusters and spatial outliers of HIV/AIDS cases 63 provinces/cities in Vietnam in 2017. The local Moran's I statistic is first employed to measure spatial auto-correlation between the number of HIV/AIDS cases in each province/city, and then identify the spatial clusters and spatial outliers of HIV/AIDS cases. Spatial distribution of HIV/AIDS clusters and outliers will be mapped with the help of a GIS. Finally, the main findings will be discussed and summarised. High numbers of HIV/AIDS cases were mainly concentrated in the provinces of the north central region, Da Nang and provinces in the south of Vietnam. Low numbers of HIV/AIDS cases were detected in the northeastern provinces, central and southeastern provinces of Vietnam. Specifically, one high-high cluster and six low-low spatial clusters, and four low-high and high-low spatial outliers of HIV/AIDS cases were successfully detected. Whereas, the only high-high spatial cluster was discovered in Binh Duong province with 3598 HIV/AIDS cases. Low-high spatial clusters were detected in four provinces including Tay Ninh (3109 cases), Long An (2799 cases), Tien Giang (2037 cases) and Ba Ria Vung Tau (2360 cases). The only high-low spatial outliers was identified in Quang Ngai (5,430 HIV/AIDS cases). It can be concluded that the local Moran's I statistic can help to effectively identify spatial clusters and spatial outlier of HIV/AIDS. Findings in this study provide an insight into how to use spatial statistics to study the spread of HIV/AIDS.

Compliance with ethical standards

Acknowledgments

The authors thank of the Vietnam Ministry of Health for providing the data and the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions

Disclosure of conflict of interest

The authors declare no conflict of interest.

Statement of informed consent

Informed consent was obtained from all individual participants included in the study.

References

- [1] Piot P, Plummer FA, Mhalu FS, Lamboray J-L, Chin J, Mann JM. AIDS: an international perspective. *Science* (80-). 1988, 239(4840):573–9.
- [2] Yuan F, Liu L, Liu L, Zeng Y, Zhang L, He F, et al. Epidemiological and spatiotemporal analyses of HIV/AIDS prevalence among older adults in Sichuan, China between 2008 and 2019: A population-based study. *Int J Infect Dis*. 2021, 105:769–75.
- [3] Bernabé E, Hay RJ, Wolfe CDA, Shibuya K. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2017, 390(10100):1260–344.
- [4] UNAIDS. Global HIV & AIDS statistics: 2020 fact sheet [Internet]. 2021. Available from: <https://www.unaids.org/en/resources/fact-sheet>
- [5] Africa SS. Statistical release: Mid-Year population estimates. 2020.
- [6] Kharsany ABM, Cawood C, Khanyile D, Lewis L, Grobler A, Puren A, et al. Community-based HIV prevalence in KwaZulu-Natal, South Africa: results of a cross-sectional household survey. *Lancet HIV*. 2018, 5(8):e427–37.
- [7] Kharsany ABM, Cawood C, Lewis L, Yende-Zuma N, Khanyile D, Puren A, et al. Trends in HIV prevention, treatment, and incidence in a hyperendemic area of KwaZulu-Natal, South Africa. *JAMA Netw Open*. 2019, 2(11):e1914378–e1914378.
- [8] Grobler A, Cawood C, Khanyile D, Puren A, Kharsany ABM. Progress of UNAIDS 90-90-90 targets in a district in KwaZulu-Natal, South Africa, with high HIV burden, in the HIPSS study: a household-based complex multilevel community survey. *Lancet HIV*. 2017, 4(11):e505–13.
- [9] Huerga H, Van Cutsem G, Ben Farhat J, Puren A, Bouhenia M, Wiesner L, et al. Progress towards the UNAIDS 90–90-90 goals by age and gender in a rural area of KwaZulu-Natal, South Africa: a household-based community cross-sectional survey. *BMC Public Health*. 2018, 18(1):1–8.
- [10] Shisana O, Rehle T, Simbayi LC, Zuma K, Jooste S, Zungu N, et al. South African national HIV prevalence, incidence and behaviour survey, 2012. 2014,
- [11] De Oliveira T, Kharsany ABM, Gräf T, Cawood C, Khanyile D, Grobler A, et al. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. *lancet HIV*. 2017, 4(1):e41–50.
- [12] Wand H, Dassaye R, Reddy T, Yssel J, Ramjee G. Geographical-level contributions of risk factors for HIV infections using generalized additive models: results from a cohort of South African women. *AIDS Care*. 2019, 31(6):714–22.
- [13] Liu H, Lin X, Xu Y, Chen S, Shi J, Morisky D. Emerging HIV epidemic among older adults in Nanning, China. *AIDS Patient Care STDS*. 2012, 26(10):565–7.
- [14] Lu L, Jia M, Ma Y, Yang L, Chen Z, Ho DD, et al. The changing face of HIV in China. *Nature*. 2008, 455(7213):609–11.

- [15] Zulu LC, Kalipeni E, Johannes E. Analyzing spatial clustering and the spatiotemporal nature and trends of HIV/AIDS prevalence using GIS: the case of Malawi, 1994-2010. *BMC Infect Dis.* 2014, 14(1):1–21.
- [16] Tanser F, Barnighausen T, Cooke GS, Newell M-L. Localized spatial clustering of HIV infections in a widely disseminated rural South African epidemic. *Int J Epidemiol.* 2009, 38(4):1008–16.
- [17] Kalipeni E, Zulu L. Using GIS to model and forecast HIV/AIDS rates in Africa, 1986–2010. *Prof Geogr.* 2008, 60(1):33–53.
- [18] Nykiforuk CIJ, Flaman LM. Geographic information systems (GIS) for health promotion and public health: a review. *Health Promot Pract.* 2011, 12(1):63–73.
- [19] Nguyen TA, Oosterhoff P, Hardon A, Tran HN, Coutinho RA, Wright P. A hidden HIV epidemic among women in Vietnam. *BMC Public Health.* 2008, 8:1–11.
- [20] Vietnam SR of. Second country report on following up the implementation to the declaration of commitment on HIV/AIDS January 2003–December 2005. MOH Hanoi, 2005.
- [21] Adams SJ. Vietnam's health care system: a macroeconomic perspective. *Citeseer*, 2005.
- [22] Tobler W. On the first law of geography: A reply. *Ann Assoc Am Geogr.* 2004, 94(2):304–10.
- [23] Anselin L. Local indicators of spatial association—LISA. *Geogr Anal.* 1995, 27(2):93–115.
- [24] Geary RC. The contiguity ratio and statistical mapping. *Inc Stat.* 1954, 5(3):115–46.
- [25] Getis A. A history of the concept of spatial autocorrelation: A geographer's perspective. *Geogr Anal.* 2008, 40(3):297–309.
- [26] Mauricio Bini L, Diniz-Filho JAF, Rangel TF, Akre TSB, Albaladejo RG, Albuquerque FS, et al. Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography (Cop).* 2009, 32(2):193–204.
- [27] Cliff AD, Ord JK. *Spatial processes: models & applications.* (No Title). 1981,
- [28] Getis A, Ord JK. *Local spatial statistics: An overview.* *Spatial analysis: Modeling in a GIS environment.* Longley, P., and M. Batty. Wiley, New York, 1996.
- [29] Vu D-T, Nguyen T-T, Hoang A-H. Spatial clustering analysis of the COVID-19 pandemic: A case study of the fourth wave in Vietnam. *Geogr Environ Sustain.* 2021, 14(4).
- [30] Nguyen TT, Vu TD. Identification of multivariate geochemical anomalies using spatial autocorrelation analysis and robust statistics. *Ore Geol Rev.* 2019, 111.
- [31] Anselin L, Syabri I, Kho Y. *GeoDa: an introduction to spatial data analysis.* In: *Handbook of applied spatial analysis: Software tools, methods and applications.* Springer, 2009. p. 73–89.