

Mapping priority neighborhoods: A novel approach to cluster identification in HIV/AIDS population

Soushieta Jagadesh (✉ soushieta.jagadesh@ird.fr)

Institut de recherche pour le developpement <https://orcid.org/0000-0003-2985-7430>

Marine Combe

Institut de recherche pour le developpement

Pierre Couppié

Centre Hospitalier Andree Rosemon

Rodolphe Elie Gozlan

Institut de recherche pour le developpement

Mathieu Nacher

Centre Hospitalier Andree Rosemon

Research article

Keywords: Human disease; disseminated histoplasmosis; urban effect; risk behavior; Self Organizing Maps

Posted Date: March 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-17843/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Urban disadvantaged neighborhoods have higher HIV risk behavior and higher levels of AIDS-related mortality. Studies demonstrate that interventions at the community level focusing on risk groups have increased success rates than individual patient-based management in the context of HIV/AIDS. We tested a novel approach to identify population groups in need of greater public health efforts to achieve UNAIDS 90-90-90.

Methods

We extracted retrospective data on 2141 HIV/AIDS patients, recruited from 1997-2017 in the regional hospitals in French Guiana. Self-organizing maps were constructed and clusters were identified based on demographic and socioeconomic variables such as age, sex, CD4 counts at Nadir, type of neighborhood, unemployment rate, and presence of opportunistic illness such as Histoplasmosis and Hepatitis B in the sample population.

Results

Neighborhood unemployment rates were identified to have a large impact in the distribution of HIV/AIDS. Also, the risk of disseminated histoplasmosis, the most common AIDS-defining illness in French Guiana, was not associated to any particular neighborhood suggesting that urban socioeconomic features are not the primary drivers of exposure risk.

Conclusion

Socioeconomically disadvantaged neighborhoods remain hotspots for HIV/AIDS. We conclude that SOM is an effective tool in the identification of risk clusters that may guide public health efforts to optimize HIV prevention and testing in French Guiana and other developing countries.

Background

The residential environment, particularly disadvantaged neighborhoods, has been shown to affect individual behavior and health through direct or indirect socioeconomic processes (1). Neighborhoods have emerged as a determinant of public health providing both objective physical infrastructure and social characteristics that influence individual and community health (2). Exploring neighborhood effects is thus important as it looks at community-based interventions to promote population health and health equity in regions with racial/ethnic minorities.

Characteristics of disadvantaged neighborhoods often correlate with individual HIV risk behavior (3). High school dropout rates, low employment rate, substance and alcohol abuse, increased street violence

usually correlate with greater HIV risk (4). Furthermore, neighborhoods with such features clustering exhibited higher HIV mortality and delayed anti-retroviral therapy (ART) initiation (3).

Histoplasmosis is one of the most frequent, often overlooked endemic disseminated mycosis in HIV patients in the Americas (5,6). The disease caused by the inhalation of the microconidia and mycelial forms of a fungus, *Histoplasma capsulatum* var. *capsulatum* (HC), found in guano-enriched soils. In French Guiana (FG), histoplasmosis is the most common AIDS-defining event and the leading cause of AIDS related deaths (7), with 75% of HIV-infected persons being foreign citizens. The fight against HIV has struggled to reach the undiagnosed reservoir with 30% of patients in Cayenne and 50% in Saint Laurent du Maroni, the two major cities of FG, diagnosed with advanced HIV-disease, with certain population groups particularly at risk of late testing. Although individual epidemiologic risk factors are important, spatialized community approaches to HIV programs seem to have an operational advantage. We thus tested the hypotheses that the distribution of HIV/AIDS and that of AIDS-related histoplasmosis varied between neighborhoods using self-organized maps, a two-dimensional data visualization tool that is trained using an unsupervised process.

Methods

Study Settings

FG is a French overseas territory located in the intertropical zone of South America. In January 2015, France rolled out “le Quartier Prioritaire de la Politique de la Ville (QPPV)” or priority district of the city’s policy to identify and improve the socially disadvantaged areas in France. Neighborhoods were identified by the poverty rate defined by INSEE as the proportion of the population living in regions that are under 60% of the median metropolitan living standards. In FG, the QPPV identified 32 priority neighborhoods in the urban, suburban, and peri-urban regions. In our study, we refer to the QPPV neighborhoods as urban disadvantaged neighborhoods.

In FG, all patients diagnosed with HIV receive free ART regardless of their socioeconomic status or country of origin. HIV care is accessible and is comparable to mainland France.

Patient Data

HIV-positive patients in follow-up at the hospitals of major cities in FG between April 1995 and May 2017 were enrolled in the French Hospital Database for HIV (FHDH). Western Blot was used to confirm the HIV status of all patients enrolled in the study.

Self-organizing maps (SOMs)

We used self-organizing maps to identify clusters who would benefit most from public health measures. SOM are a type of artificial neural network that consists of an array of units, called nodes, arranged in a fixed position on a grid. The key feature of the SOM is that the topology of the original input data is conserved, i.e. similar variables are grouped together on the map. We trained the SOM model of 2120

nodes and variables ($n = 9$) with 2000 iterations at a learning rate between 0.05–0.01 to reach a minimum plateau. The nodes were modeled on a 20×20 grid with a hexagonal topology. Each node represented an HIV patient and his/her neighborhood with its position fixed in the grid. The variables included age, sex, CD4 counts at Nadir, neighborhood type, CDC classification of HIV, histoplasmosis/hepatitis positivity, unemployment rate and neighborhood security. All SOM statistics were performed using the 'kohonen' package in R version 3.6.1.

Ethics

All patients enrolled in the FHDH gave written informed consent to the use of data for research. The data is anonymized and encrypted before transfer to the Ministry of health and the Institut National de la Recherche Médicale (INSERM), which centralize data from Regional Coordination for the fight against HIV (COREVIH) across France. This cohort has been ethically approved by the Commission Nationale Informatique et Libertés (CNIL) since 1992 and has been used for numerous publications.

Results

Sociodemographic characteristics

We recruited a sample population of 2141 patients diagnosed with HIV between 1997 and 2017 in FG. The sample consisted of 1101 (51.4%) male and 1040 (48.6) female patients with an average age of 37.4 ± 0.26 years at the time of diagnosis. The patients were classified based on CDC classification system of HIV infection; A ($n = 1256$ [58.7%]), B ($n = 249$ [11.6%]), and C ($n = 636$ [29.7%]) and had an average CD4 count of 205.7 cells/mm³. Our dataset included cases from rural regions ($n = 356$ [16.6%]) and socioeconomically disadvantaged ($n = 1455$ [68%]) and favorable ($n = 330$ [15.4%]) neighborhoods. The HIV patients from disadvantaged neighborhoods had 1.33 times [95% CI: 1.01–1.75; $P 0.04$] higher odds of progressing to AIDS in comparison to the favorable ones. We found no significant risk of developing histoplasmosis [0.66–1.82; 0.74] in HIV patients inhabiting poorer neighborhoods.

SOM model and clusters

The heatmaps illustrate the distribution of seven variables in the sample population (Fig. 1A-G). We used hierarchical agglomerative clustering (HAC) to determine the optimal number of clusters in the sample population (Fig. 2). The demographic and socio-economic clusters in the HIV-related histoplasmosis population in FG is visualized in Fig. 1H. We observed that the largest cluster (Fig. 1H; blue; $n = 1365$) included PHLA) with a low employment rate living in urban disadvantaged neighborhoods and rural areas. The cluster also included regions where crime and insecurity are higher than in richer neighborhoods. The significantly smaller second cluster (Fig. 1H; green,) was represented by cases from favorable suburbs with a higher employment rate and relatively higher proportion of cases belonging to CDC class A. The histoplasmosis (Fig. 1H; purple; $n = 113$) and hepatitis (red; $n = 141$) cases were grouped in separate clusters. The hepatitis and histoplasmosis clusters had significantly higher males ($P < 0.0001$), increased mortality ($P = 0.009$, $P = 0.002$), and lower CD4 counts at nadir ($P = 0.008$, $P < 0.0001$),

respectively. The descriptive statistics are detailed in Supp.Table.1. The smallest cluster (Fig. 1H; orange; n = 109) were cases with higher CD4 counts at nadir.

Discussion

Here, for the first time, we mapped the neighborhood effect on HIV/AIDS using self-organizing maps and showed that the spatial boundaries of HIV/AIDS distribution follow those of neighborhoods. Our second hypothesis that HIV-related histoplasmosis risk differed by neighborhood was rejected as the risk of developing histoplasmosis in HIV/AIDS was similar in both the socioeconomically privileged and underprivileged neighborhoods.

We found unemployment to have a higher structural impact than the other tested variables on the identified clusters. Studies demonstrate that unemployed HIV patients had an increased risk of HIV mortality and disease progression than those with stable employment in the HAART era (8). This coincides with our results, which showed that HIV-infected persons in economically poorer neighborhoods had higher odds of progressing to AIDS. Unemployed individuals have lower access to - and/or- underutilize preventive healthcare services than their employed counterparts (9), which potentially delays HIV diagnosis and treatment. Furthermore, economic insecurity amongst youth, particularly women, has been shown to increase the HIV burden in developing countries including FG (10,11).

Our results showed that neighborhood insecurity formed a sizable portion of the largest cluster. Research has shown that neighborhood crime was associated with HIV risk taking behaviors such as unprotected sex and multiple sexual partnerships (11,12). Furthermore, a recent study observed that HIV-related mortality reached over 8% of released prison inmates in FG (13).

UNAIDS aims for 90% of HIV-infected persons aware of their diagnosis, 90% of those diagnosed on antiretroviral treatment and 90% of these virologically suppressed by 2020. However, the first 90% seems hard to reach in most countries and notably in FG where the proportion of patients diagnosed each year with advanced HIV disease remains stable despite efforts to scale up and diversify HIV-testing. The present approach may bring strategic and operational insights to improve the capacity to reach the hidden reservoir of undiagnosed infections (14).

We observed that the risk of histoplasmosis in PLHA was similar in urban neighborhoods and rural regions. Thus, it seemed that the incidence of histoplasmosis in FG was ubiquitous, influenced by the environmental distribution and endemicity of spores rather than socioeconomic factors or microenvironmental nuances. In our study, histoplasmosis, the most frequent AIDS-defining infection in FG, was a significant cluster in PLHA (15). Our results on the histoplasmosis cluster is consistent with previous studies from FG (7), where the risk factors include male sex and < 50 cells/mm³ CD4 at nadir. We also observed higher mortality in histoplasmosis and hepatitis clusters, which demonstrates the significance of cluster identification for community-level management of HIV/AIDS.

We did not analyze the environmental factors influencing the distribution of histoplasmosis, providing thus scope for future studies to analyze the impact of soil acidity and rainfall on disease incidence. In addition, the lack of socioeconomic variables, such as built environment, is another limitation of our study.

Conclusions

We conclude that SOMs applied to HIV/AIDS cases are an effective tool to identify and prioritize clusters in a large dataset. Whilst unemployment and neighborhood insecurity have a significant impact on HIV risk and socioeconomically disadvantaged neighborhoods remain hotspots for HIV/AIDS, histoplasmosis was found dependent on the environmental distribution of the pathogen's spores. Public health efforts should prioritize disadvantaged neighborhoods for HIV/STI awareness, and screening and management of hepatitis and histoplasmosis.

Abbreviations

AIDS
Acquired Immunodeficiency Syndrome
ART
Anti-Retroviral Therapy
CDC
Center for Disease Control
FG
French Guiana

Declarations

Ethics approval and Consent to participate

All patients enrolled in the FHDH gave written informed consent to the use of data for research. The data is anonymized and encrypted before transfer to the Ministry of health and the Institut National de la Recherche Médicale (INSERM), which centralize data from Regional Coordination for the fight against HIV (COREVIH) across France. This cohort has been ethically approved by the Commission Nationale Informatique et Libertés (CNIL), the regulatory authority, in 1992 and has been used for numerous publications.

Consent for Publication

Not applicable

Competing interests

All authors declare no conflict in interests.

Author's contributions

SJ conceived the study design, analysed the data and carried out the statistical analyses. Both MC and PC contributed to the final version of the manuscript. REG and MN helped to supervise the project. SJ, REG and MN wrote the paper. All authors have read and approved of the manuscript.

Acknowledgements

None

Funding

Not applicable

Availability of data and materials

The datasets generated and/or analyzed during the current study are not publicly available due to the sensitive information including potential geographic identifiers of the cases of a disease, HIV/AIDS, which remains stigmatized in the region. However, are available from the corresponding author on reasonable request.

References

1. Mayer SE, Jencks C. Growing up in Poor Neighborhoods: How Much Does it Matter?. Vol. 243, Science. American Association for the Advancement of Science; [cited 2019 Nov 8]. p. 1441–5. Available from: <https://www.jstor.org/stable/1703125>
2. Koh HK, Oppenheimer SC, Massin-Short SB, Emmons KM, Geller AC, Viswanath K. Translating research evidence into practice to reduce health disparities. *Am J Public Health*. 2010 Apr 1 [cited 2019 Nov 8];100 Suppl:S72-80. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2837437&tool=pmcentrez&rendertype=abstract>
3. Latkin CA, German D, Vlahov D, Galea S. Neighborhoods and HIV: A social ecological approach to prevention and care. *Am Psychol*. 2013;68(4):210–24.
4. Singer MC, Erickson PI, Badiane L, Diaz R, Ortiz D, Abraham T, et al. Syndemics, sex and the city: Understanding sexually transmitted diseases in social and cultural context. *Soc Sci Med*. 2006 Oct;63(8):2010–21.
5. Nacher M, Adenis A, Arathoon E, Samayoa B, Lau-Bonilla D, Gomez BL, et al. Disseminated histoplasmosis in Central and South America, the invisible elephant: The lethal blind spot of international health organizations. Vol. 30, AIDS. Lippincott Williams and Wilkins; 2016. p. 167–70.
6. Nacher M, Leitao TS, Gómez BL, Couppié P, Adenis A, Damasceno L, et al. The Fight against HIV-Associated Disseminated Histoplasmosis in the Americas: Unfolding the Different Stories of Four

- Centers. *J Fungi* [Internet]. 2019 Jun 17 [cited 2019 Nov 12];5(2):51. Available from: <https://www.mdpi.com/2309-608X/5/2/51>
7. Nacher M, Adenis A, Blanchet D, Vantilcke V, Demar M, Basurko C, et al. Risk Factors for Disseminated Histoplasmosis in a Cohort of HIV-Infected Patients in French Guiana. *PLoS Negl Trop Dis*. 2014;8(1):7.
 8. Maruthappu M, Zhou C, Williams C, Zeltner T, Atun R. Unemployment and HIV mortality in the countries of the Organisation for Economic Co-operation and Development: 1981–2009. *JRSM Open*. 2017 Jul;8(7):205427041668520.
 9. Tefft N, Kageleiry A. State-level unemployment and the utilization of preventive medical services. *Health Serv Res*. 2014 Feb;49(1):186–205.
 10. Austin KF, Choi MM, Berndt V. Trading sex for security: Unemployment and the unequal HIV burden among young women in developing nations. *Int Sociol*. 2017 [cited 2019 Nov 27];32(3):343–68. Available from: <https://doi.org/10.1177/0268580917693172>
 11. Nacher M, Vantilcke V, Parriault MC, Van Melle A, Hanf M, Labadie G, et al. What is driving the HIV epidemic in French Guiana? *Int J STD AIDS*. 2010 May;21(5):359–61.
 12. Ojikutu BO, Bogart LM, Klein DJ, Galvan FH, Wagner GJ. Neighborhood crime and sexual transmission risk behavior among black men living with HIV. *J Health Care Poor Underserved*. 2018 Feb 1;29(1):383–99.
 13. Huber F, Merceron A, Madec Y, Gadio G, About V, Pastre A, et al. High mortality among male HIV-infected patients after prison release: ART is not enough after incarceration with HIV. 2017 [cited 2019 Dec 4]; Available from: <https://doi.org/10.1371/journal.pone.0175740>
 14. Nacher M, Huber F, Adriouch L, Djossou F, Adenis A, Couppié P. Temporal trend of the proportion of patients presenting with advanced HIV in French Guiana: Stuck on the asymptote? *BMC Res Notes*. 2018 Nov 26;11(1).
 15. Nacher M, Adenis A, Adriouch L, Dufour J, Papot E, Hanf M, et al. Short report: What is AIDS in the Amazon and the Guianas? Establishing the burden of disseminated histoplasmosis. *Am J Trop Med Hyg*. 2011 Feb;84(2):239–40.

Figures

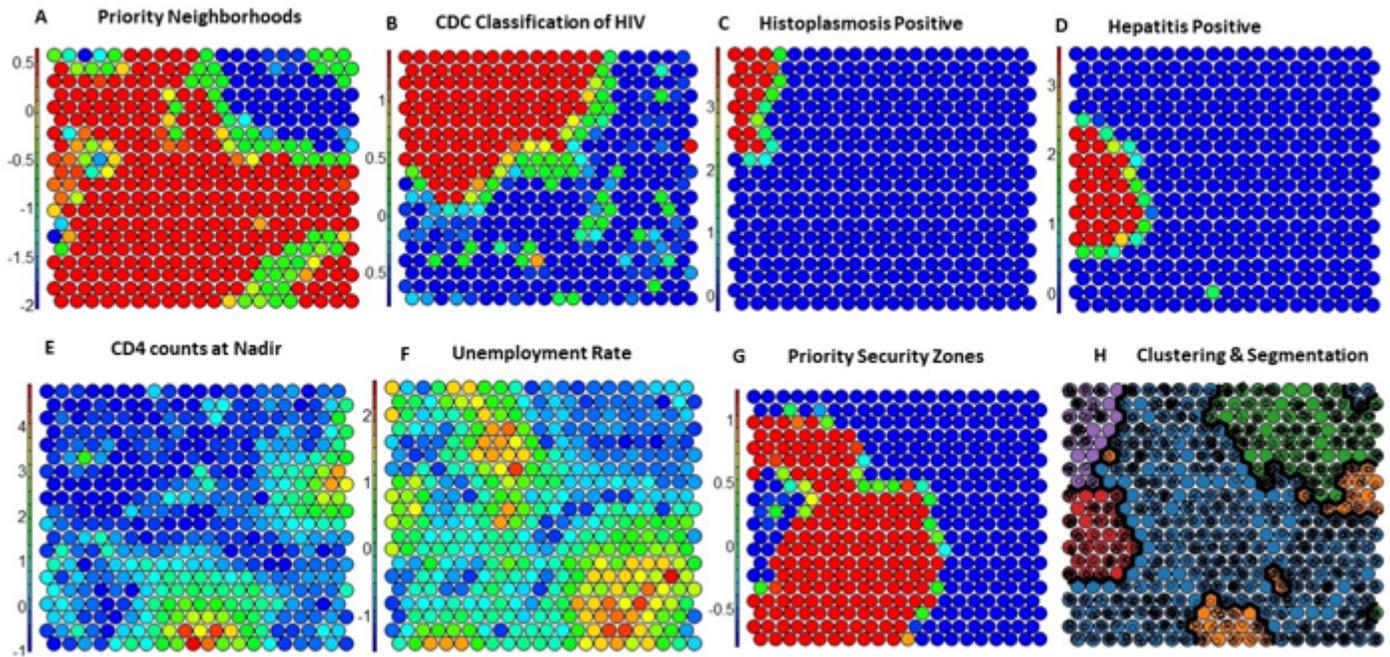


Figure 1

Self-organizing maps (SOM) model of 2120 nodes on a 20x20 grid. The spectrum of colors represented in the y-axis represents the scale of the map. A) demonstrates the structure of HIV/AIDS population where the patient living in urban disadvantaged neighborhoods are illustrated (in red), in rural (yellow-green spectrum) and favorable urban neighborhoods (blue); B-E) illustrate the spectrum of severity of HIV/AIDS; F & G) illustrate the neighborhood unemployment rate and insecurity; H) represents the hierarchical agglomerative clustering and segmentation on the SOM model.

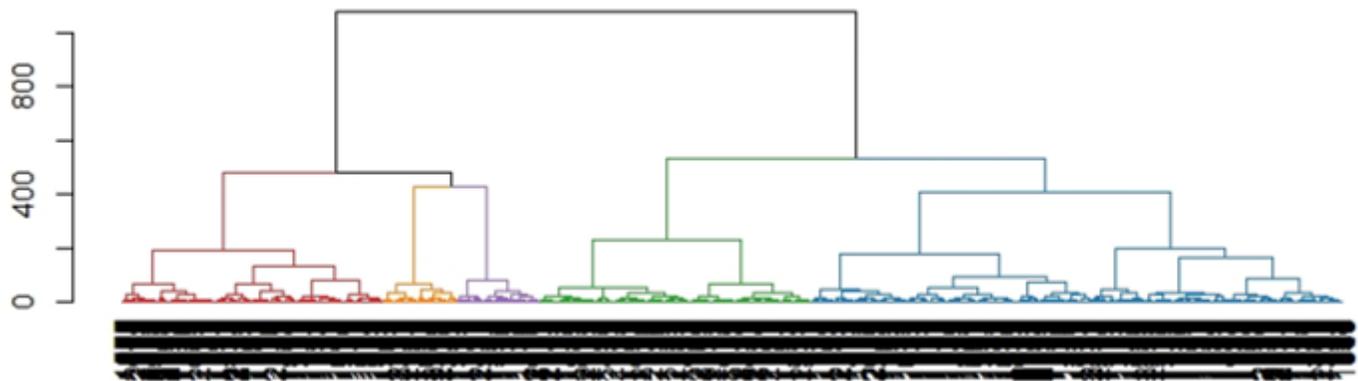


Figure 2

Cluster dendrogram to determine the optimal number of clusters in the SOM model. The colors correspond to the colors of the SOM model.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterialaids.docx](#)