Geographical models in epidemiology

Diego F. Cuadros
Assistant Professor
Department of Geography and GIS
University of Cincinnati

June 20, 2019
Cholera outbreak

London 1854
Cholera outbreak

- Cholera epidemic in London 1854

- **John Snow** used a dot map to illustrate the cluster of cholera cases around the **water pump** to show the connection between the quality of the water source and cholera cases.

- Snow's study was a major event in the history of public health and geography. It is regarded as the **founding event** of the science of **epidemiology**.
What is Health Geography?

• Health geography is the application of geographical perspectives and methods to the study of health, disease and health care

• Health geography uses the concepts and techniques of geography to investigate health-related issues

• Health geography, uses the tools and approaches of geography to tackle health-related questions focused on the importance of variations across space, with an emphasis on concepts such as location, direction, and place
Malaria
Schistosomiasis
Dengue Fever
Zika virus

Environmental suitability for Zika virus

1
0
Ebola virus

[Map of Africa showing Ebola outbreaks with data on infections and deaths.]

Sources: WHO; IUCN

Economist.com/graphicdetail
Health geography and spatial epidemiology has gained popularity and relevance during the recent years, why?
Spatial data

Malaria sample sites

Demographic Health surveys sample sites
Social networks

The image is of London; red dots represent Instagram uploads and blue dots are the tweets. The white dots show us which locations post to both.
Cholera outbreak
Disease mapping examples
The **Malaria Atlas Project** sheds new light on the impact of malaria control in Africa.

Malaria in Africa has halved since the turn of the millennium; 663 million cases have been averted.
Ecological Niche Modeling: Vector-borne diseases

Cofactors

Environmental variables:

- Normalized Difference Vegetation Index (NDVI)
- Temperature
- Precipitation
- Evapotranspiration
- Elevation
Ecological Niche Modeling

Best Approximating Model(s)

RSF probability map

Selected: used to degree exceeding availability

Avoided: used less than expected from availability

Human-Use Variables

Landscape Variables

Vegetation Variables

Use / Avail. data

Resources

Telemetry (red), Random (yellow) points
Habitat Evaluation Procedures

Logistic regression:

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 = \text{logit}(p) \]

\[ \Pr(Y = 1 \mid \text{the explanatory variables } x) = \pi \]

\[ \pi = e^{-\text{logit}(p)} / [1 + e^{-\text{logit}(p)}] \]
High resolution maps of the **geographic distribution of a disease**. The use of the survey data from a **sample of locations** to predict continuous surfaces of risk, informed by **environmental and demographic covariates**.
Malaria
Spatio-temporal dynamics of malaria
Introduction: HIV

• The human immunodeficiency virus (HIV) is retrovirus that causes HIV infection and over time acquired immunodeficiency syndrome (AIDS)

• AIDS is a condition in humans in which progressive failure of the immune system allows life-threatening opportunistic infections to thrive

• Without treatment, average survival time after infection with HIV is estimated to be 9 to 11 years

• Infection with HIV occurs by the transfer of blood, semen, vaginal fluid, or breast milk
Introduction: HIV

Vector-borne diseases

Sexually transmitted infections (STI)

Pathogen → Vector → Host

Infected → Susceptible
Is the spatial distribution of HIV random? Clustered?
For each country, we only considered the most recent Demographic Health Survey where HIV data were collected. As a result, a total of **20 countries in SSA were included**
Methods: Clustering analysis

- **Spatial scan statistics** is a cluster detection test able to find the location of areas with **higher or lower numbers of cases** (for instance HIV infections) than expected.

- For each potential cluster, a likelihood ratio test was computed. The numbers of observed and expected HIV infections within and outside the circular window were then compared to test the **null hypothesis of spatial randomness**.
• We identified 38 clusters with high HIV prevalence (hot spots; red circles), and 45 clusters with low HIV prevalence (cold spots; blue circles)

• The HIV seroprevalence within the clusters with high HIV seroprevalence ranged from 1.9% in a cluster in Senegal to 30.8% in a cluster in Zimbabwe, with a median of 11.5%
The relative risk of HIV infection for individuals within clusters with high HIV seroprevalence was **negatively associated with HIV prevalence** of the corresponding country (p < 0.001)

The results of our analysis indicate **stark geographical variation in HIV prevalence** in most of the countries. The observed spatial variation in HIV prevalence highlights **a clustered HIV transmission across SSA** within micro-epidemics of different scales.
Could environmental and socio-economic and behavioral factors be used to generate high resolution maps of HIV prevalence in SSA?
Ecological Niche Modeling

Best Approximating Model(s)

RSF probability map

Selected: used to degree exceeding availability

Avoided: used less than expected from availability

Human-Use Variables

Landscape Variables

Vegetation Variables

Use / Avail. data

Resources

Telemetry (red), Random (yellow) points
HIV factors

Four countries
Tanzania, Kenya, Malawi and Mozambique

Environmental (geographical) factors:
• Normalized Difference Vegetation Index (NDVI)
• Population
• Distance to main roads

Socio-economic and behavioral factors:
• Wealth index
• Male circumcision
• Lifetime sexual partners
• Education
• Ever been tested for HIV
• Condom use
Maps of cofactors

Maps of cofactors

+ + + +

= High Resolution HIV prevalence map
High resolution map of HIV in Tanzania
High resolution maps for HIV prevalence in East Africa

High resolution maps for HIV prevalence in (A) Kenya; (B) Malawi; (C) Mozambique; and (D) Tanzania.

HIV prevalence distribution in Africa


Absolute change in HIV prevalence among adults aged 15–49 between 2000 and 2017 at the country level (a), first administrative subdivision level (b), second administrative subdivision level (c) and 5 × 5-km grid-cell level (d).
HIV prevalence distribution in Africa

AHRI Surveillance System
The site has collected socio-demographic information on a population of 87,000 participants within a circumscribed geographic area (438 km$^2$) for over a decade.
The HIV hotspot contained 40.8% of the total HIV-positive individuals, and individuals located within the geographical cluster had 46% higher risk of HIV infection compared to individuals located outside the cluster.
Estimated HIV seroconversions per year were highly concentrated within the HIV hotspot with an average of 0.04 seroconversions per year per 100 m², compared to 0.01 seroconversions per year per 100 m² in the area outside of the HIV hot-spot.
Why mapping diseases?


- The identification of the settings where both the burden of disease and the drivers of the disease are concentrated could play an important role for optimization of resource allocation based on geographically targeted interventions.
Resource allocation

Difference in health gains between the uniform and focused approaches
Anderson et al. *Lancet HIV* 2014
HIV prevention and care in Zimbabwe

High resolution maps of HIV prevalence in Zimbabwe for (A) Females and (D) Males; geographic dispersion of HIV-infected (B) females and (E) males in Zimbabwe. High HIV burden areas are illustrated in red for (C) females and (F) males.

Province-level maps of (A) incidence of HIV infection in 2015, (B) ART coverage, and (C) viral load suppression of those on ART. Map in (D) presents estimates of the average number of clients per ART site per district. Geographical dispersion of males (E) and females (F) lacking HIV treatment and care.

Use of mosquito bed nets in Central and East Africa

- Average of ownership of mosquito bed nets 68%
- Average use of mosquito bed nets 46%

Interpolated surface map of lower (<60%) and higher levels for use of mosquito net (A), areas with lower (Parasite Rate, PR<0.1) and higher endemic malaria (B). the combination of both A and B to identify high-risk areas (low mosquito net use and high malaria, C), and the estimated number of children at risk (D).
Human activities such as deforestation linked to agriculture, livestock production and mining activities were identified as key drivers of the spatial distribution of the cutaneous leishmaniasis epidemic in Colombia.

• Other uses: identification of vulnerable populations at high risk of infection

Spatial risk of HIV infection: Spatial structure of the transmission network
Concentrated epidemic

- General population ($R_0 < 1$)
- Behavioral high-risk subpopulation ($R_0 > 1$)
The study of sexually transmitted infections such as HIV has focused on social space.

**Sexual networks**: groups of persons connected to one another sexually.

**Characteristics**:
- Number of partners (links)
- Serial monogamy
- Concurrent relationships

**Core groups**: members that have high levels of risk behavior and can fuel sustained transmission.
Concentrated epidemic

High-risk groups
Generalized epidemic

General population
Geographical space

- HIV ‘hot-spots’ can behave as the highly connected nodes of in the transmission network
Spatial risk of HIV infection

HIV Prevalence

- > 10%
- < 2%
The spatial connectivity of the transmission network of an entire community has never been studied before, and the contribution of geographical clusters of HIV infections, or ‘hot-spots’ on the spread of the infection in the entire population is virtually unknown.

What is the contribution of the HIV hotspot in the transmission network?
Transmission links

- We geo-located and genetically sequenced 1,222 HIV-positive individuals, from whom phylogenetic transmission clusters were identified.

- We constructed the spatially explicit transmission network with 350 transmission links identified.
• 72% of the links included at least one individual located within the HIV hot-spot

• 28% of the links included individuals located outside the HIV hot-spot
• We found that more than 70% of the HIV transmission links identified were directly connected to an HIV hot-spot, illustrating the **high connectivity between the hot-spots and the general community**

• The HIV hotspots might play a key role in the HIV transmission network and could substantially contribute to the dispersion of the infection

• We hypothesize that HIV hot-spots behave as core groups in the transmission network, and **interventions targeting these hot-spots could not only reduce the levels of new infections in these geographical core groups, but also disrupt the transmission of the infection in the entire community**
Combination prevention
Combination prevention
• Other uses: disease interactions

Spatial analysis and disease mapping are powerful tools for revealing the interactions among diseases and the geographical locations where several diseases collide, thereby exposing the structure of the disease community in these areas.
Diabetes and tuberculosis in India

Clusters and interpolated surface prevalence levels (%) for self-reported diabetes in India (A). Clusters and interpolated surface exposure level (cases per 100 000 inhabitants) for reported cases of tuberculosis (TB) India (B)
Depression and tuberculosis in South Africa

- Individuals who reported being infected with TB had significantly higher odds of belonging to a hotspot of new cases of depression.

- In A) Continuous surface map of prevalence of existing cases of depression and the location of spatio-temporal clusters of existing cases of depression; B) continuous surface map of prevalence of new cases of depression and the location of spatial clusters of new cases of depression; C) Continuous surface map of prevalence of tuberculosis (TB) and the location of spatio-temporal clusters of existing TB cases. D) Geographical overlap between clusters of new cases of depression (blue circles), and clusters of TB (red circles).

Opioid overdose epidemic in the US
Social determinants of addiction

Mentally Unhealthy Days

Drug Poisoning Mortality Rate

Legend
Mentally Unhealthy Days
- 0 - 1.4
- 1.5 - 3.3
- 3.4 - 4.3
- 4.4 - 5.5
- 5.6 - 7.9

Legend
Drug Poisoning Mortality Rate
- 5.0 - 11.1
- 11.2 - 15.1
- 15.2 - 19.1
- 19.2 - 25.1
- 25.2 - 31.0
Social determinants of addiction
Some other uses

- Maps can show disparity in health care delivery
- Help patients find quality health care
- Help authorities assess risk
Maps: Great Communicators!

- Graphic representations of data frequently more powerful than numeric or textual representations.

- Maps seem authoritative to policy makers and stakeholders.
Acknowledgements

Frank Tanser

Tulio de Oliveira

Hae-Young Kim

Alain Vandormael

Andrew Tomita
Thank you!