MODELING HIV TESTING AND PREVENTION STRATEGIES - P-W1
Poster # 1154

25th CROI
Conference on Retroviruses and Opportunistic Infections

CFAR SAN DIEGO | CENTER FOR AIDS RESEARCH

FINE TUNING SPATIAL RESOLUTION OF HIV EPIDEMIOLOGIC DATA WHILE PROTECTING PRIVACY

Marta M. Jankowska[1], Tommi Gaines[2], Sanjay R. Mehta[2,3], Susan J. Little[2], Antoine Chaillon[2]

1-Qualcomm Institute, California Institute for Telecommunications and Information Technology, University of California San Diego, 2-Division of Infectious Diseases and Global Public Health, University of California San Diego, 3-Medical Service, San Diego Veterans Affairs Medical Center, CA USA.

Antoine Chaillon, MD, PhD
University of California, San Diego
9500 Gilman Drive
San Diego, CA 92093, USA
E-mail: achaillon@ucsd.edu

Background

- Advances in spatial analytics, health geography, and spatial epidemiology have demonstrated that precise measures of locations of individuals can lead to novel research findings and intervention opportunities.[1]
- Coupled with health data, geolocation is protected under the Health Insurance Portability and Accountability Act of 1996 (HIPAA). Data are often aggregated into larger spatial units to preserve privacy, limiting utility of spatial analysis such as measuring access, modeling disease transmission, clustering analysis, and optimally targeted interventions.
- Dasymetric mapping (DASY) can be coupled with small area estimation methods to generate a detailed model of the distribution of a HIV cohort, and result in fine scale spatial analytics.

Objective

- To develop a cartographic dasymetric mapping approach coupled with small area probabilistic reweighting[2] to identify clusters of new HIV infections in San Diego County.

Methods

- Dasymetric mapping (DASY) approach:
  - DASY is a type of areal interpolation where the region being mapped is divided into homogeneous target polygons with the purpose of best portraying the underlying statistical surface of the data being mapped.[3]
  - Ancillary land use data are used in DASY mapping to provide an estimate of the underlying surface.
  - Refined constraining data to small areas (i.e. microdata) are used to downscale the spatial distribution of the original dataset.
- Data: We obtained age, sex and general location data (Health and Human Services Agency (HHSA) regions) from the 657 HIV individuals enrolled in the San Diego Primary Resource Consortium (SDPIRC).
- Application (**Figure**): we present a *DASY method* with residential land use data as ancillary data (step 1) to map the underlying distribution of the SDPIRC HIV population at a 500 x 500m grid cell level. Constraining variables were modeled for each grid cell to generate a microdata set at the 500 x 500m grid cell level, including (1) age, (2) sex and (3) HIV prevalence by age and (4) by sex (step 2), resulting in the dasymetric mapping of the SDPIRC cohort (step 3).

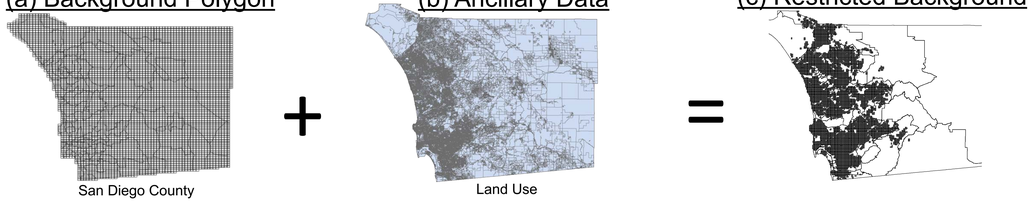
References. **1** - Cromley EK, McLafferty SL. GIS and public health. Guilford Press; 2011; **2** - Ruther M, *et al.* Deriving Small Area Mortality Estimates Using a Probabilistic Reweighting Method. Ann Am Assoc Geogr. 2017;:1–16; **3** - Eicher CL, Brewer CA. Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation. Cartogr Geogr Inf Sci. 2001;28:125–38.

Acknowledgments. This work was supported primarily by a grant from the National Institutes of Health K01DA034523 and from the Center for AIDS Research CFAR P30 AI036214. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the poster.

Results

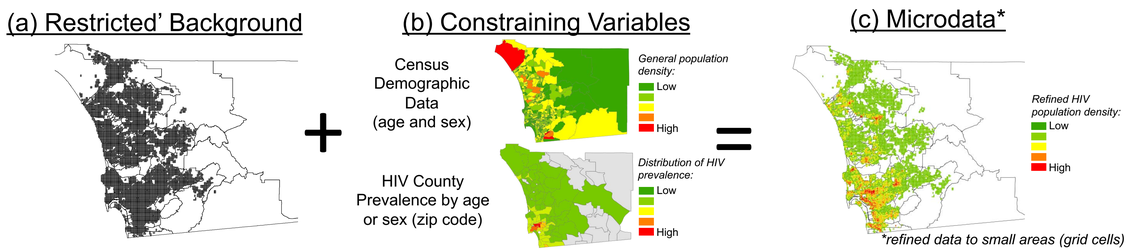
Step 1. Creating a Background Restricted Map

Step 1: Generation of a background 500x500m grid surface (a) combined with residential land use data (b). The resulting restricted map (c) is used for further areal interpolation of census and county data (step 2).



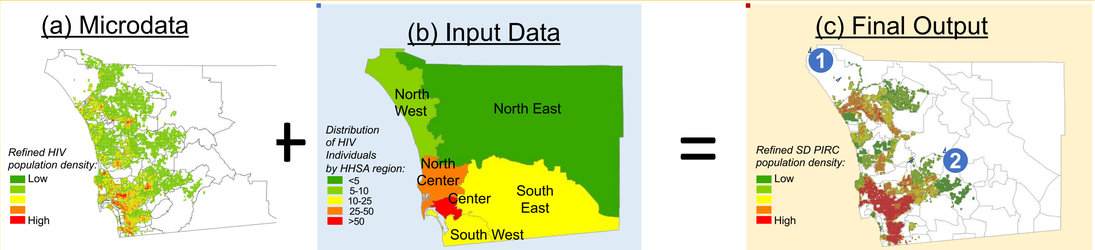
Step 2. Creating a Microdata Set Based on Publicly Available Data

Step 2: Demographic and HIV prevalence data (b) are redistributed according to land use area into the 500m grids (a) resulting in a microdata set (c). The resulting microdata is used for reweighting of the SDPIRC HIV dataset from course HHSA geographies to the finer 500m grid (step 3).



Step 3. Probabilistic Reweighting of Input Data Uncover Increased Resolution of Hotspots of HIV New Infections

Step 3: SDPIRC characteristics (b) are matched to the relative presence of the same characteristics in the microdata (a) of each grid cell, generating a likelihood map of how the SDPIRC cohort is distributed (c). Blue circles indicate local hotspots of HIV not previously identified in the HSSA map (middle panel, input data).



Conclusions

- Identifying the spread of transmissible diseases, including HIV, is crucial to optimally target treatment and prevention programs.
- Improving the granularity of health data without violating privacy and confidentiality restrictions can help to reveal spatial patterns that are not apparent in aggregated data sets.