

Mobile Technology to Measure Activities Related to Cognitive Health in Older adults



Atif Adam, PhD (1); Kyle Moored, PhD(c) (1); Breanna Crane MS (1); Thomas Chan, PhD (1); Michelle Carlson, PhD (1)

(1)Department of Mental Health, Johns Hopkins University Bloomberg School of Public Health;

Overview

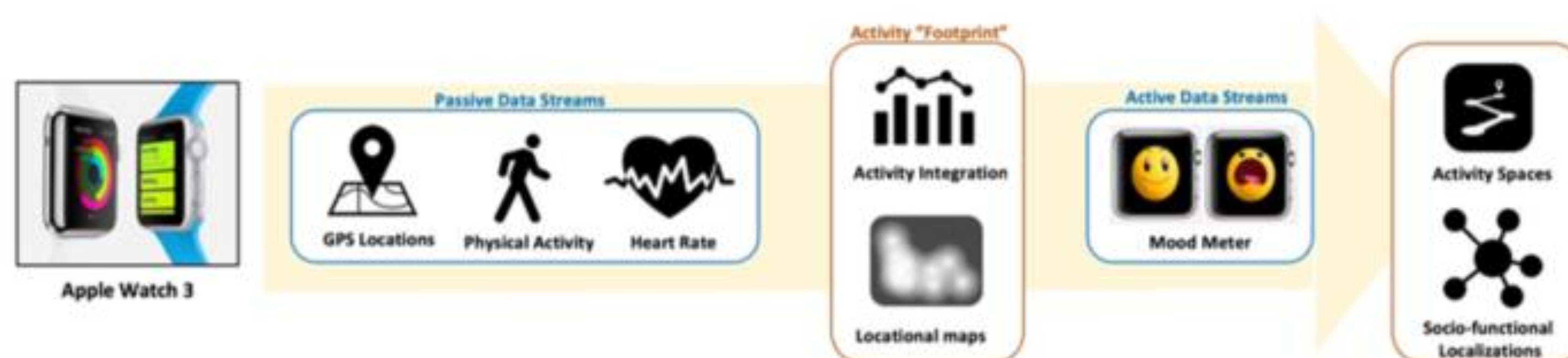
Evidence linking cognitively enriching physical activity to risk for dementia has been largely restricted to exercise, a behavior that is difficult to sustain in aging adults. This is particularly so among urban-dwelling older adults who may live in unsafe and under-resourced neighborhoods. Such limitations have left an important gap in understanding how to increase levels of physical activity in daily life and how much is sufficient to engender use-induced brain plasticity in networks implicated in dementia. A better understanding of the benefits of physical activity will require improved metrics to define patterns of activity in daily life.

Purpose

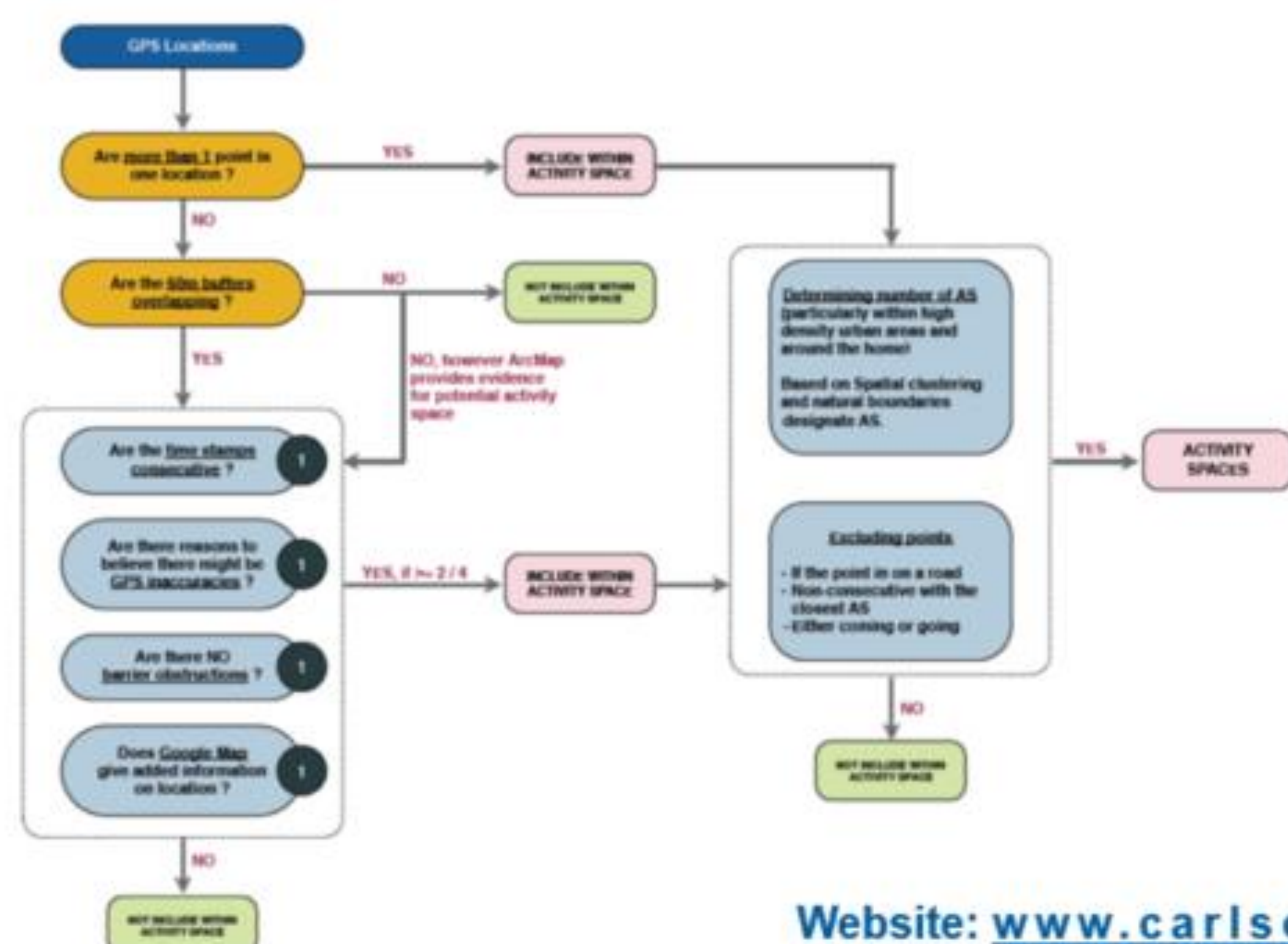
To explore the use of integrated GIS, GPS physical, accelerometer and heart rate data, to develop novel activity metrics,

Just-in-Time Adaptive Assessment Pipeline

- An emerging mobile phone intervention/assessment design.
- Adapting to the dynamics of an individual's emotional, social, physical and contextual state.
- Study design aiming to address the dynamically changing needs of individuals via the provision of the type/amount of support needed, at the right time, and only when needed.



Activity Space Algorithms



Website: www.carlsonlab.org

Results

Increasing Spatial, Temporal and Data Complexity



Conclusion

- Activity in social contexts may provide neurocognitive benefits as great as that for physical activity alone
 - 100 steps indoors may not be equivalent to 100 steps outdoors in spatially & socially complex community spaces
- Activity with purpose and makes you feel good may confer neurocognitive benefits that augment those related to exercise
 - Sustain movement in pursuit of a goal (& the less spandex involved, the better!)

Geographic Mapping of Cholera Hotspots in the East and Southeast African Countries

Mohammad Ali, Godfrey Bwire, Fausta Moshia, Ahmed Abade, Waqo Boru, John Mwaba, Allison Shaffer, Patrick Shea, Amanda Debes, Roma Chilengi, David A Sack

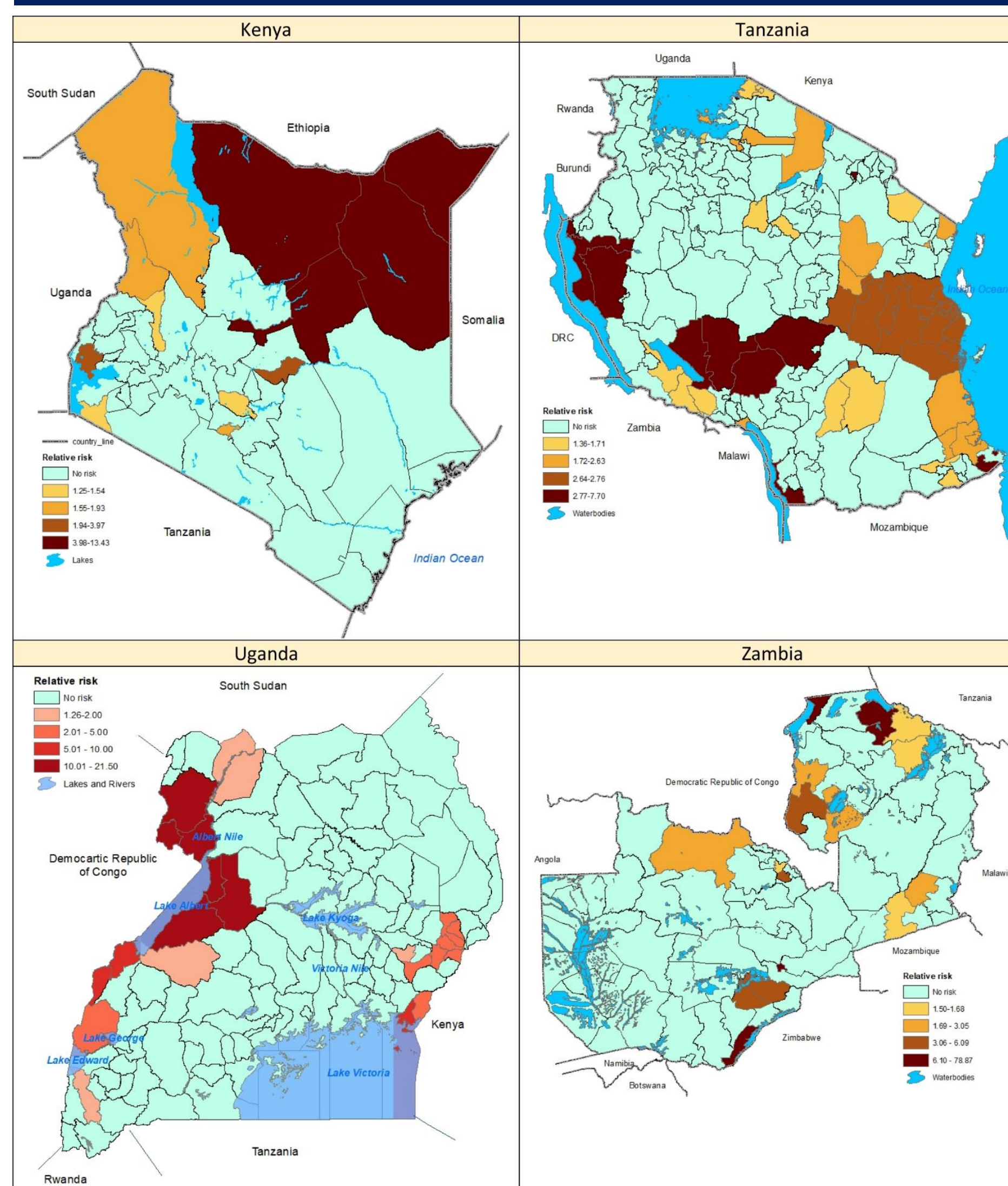
BACKGROUND

- The goal of studying cholera is its elimination from the globe
- Even with an increasing understanding, cholera outbreaks continue, appears to be increasing in many countries in sub-Saharan Africa
- WHO described a roadmap to eliminate the threat of cholera by 2030
- Critical to the success of the roadmap is the proper identification of areas of higher risk which are called the "cholera hotspots"

OBJECTIVES

To identify cholera hotspots in four East and Southeast African countries which can serve as a guide for the development of a roadmap for elimination of cholera in this part of Africa

CHOLERA HOTSPOTS

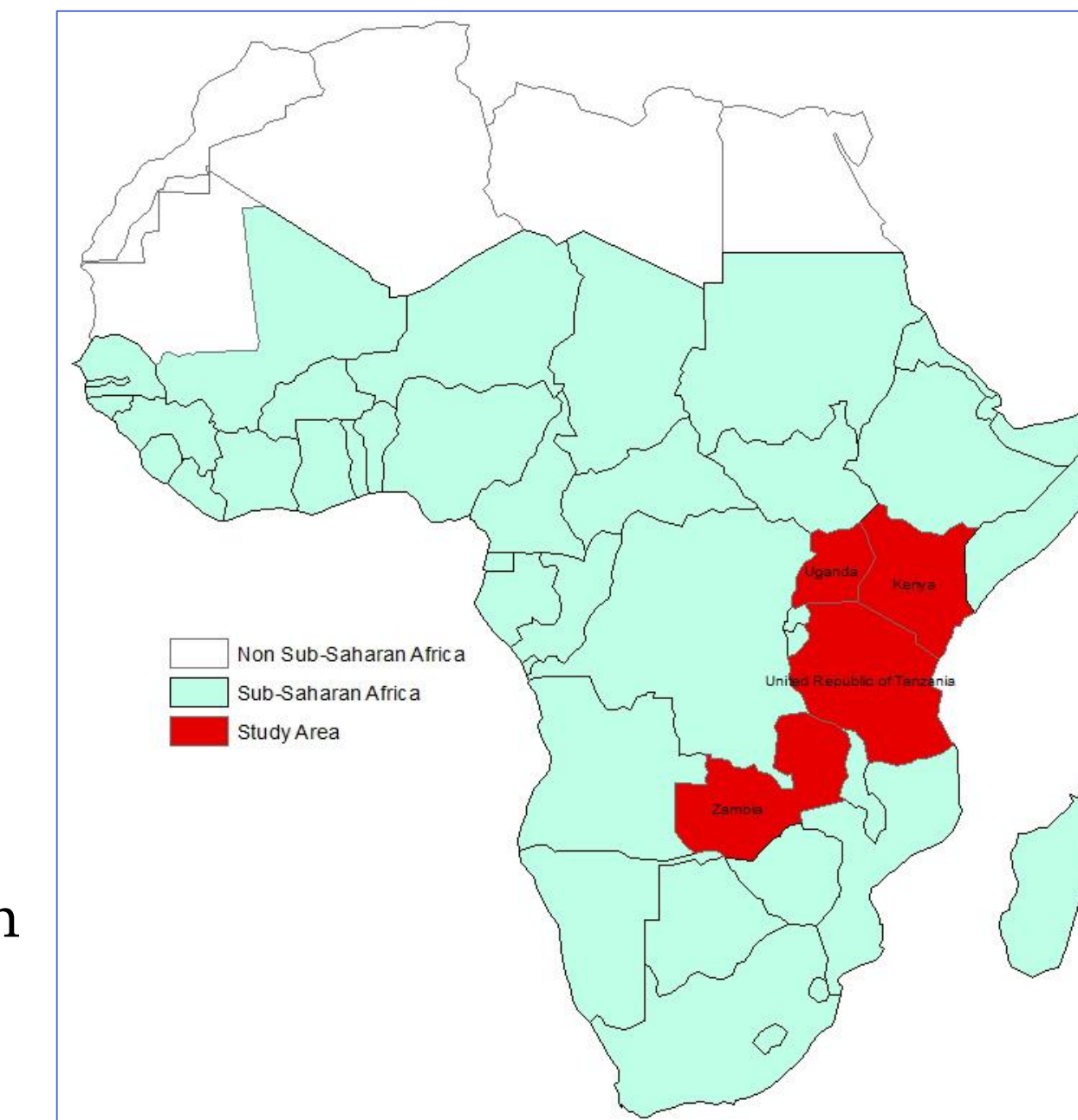


LIMITATIONS

- Included cases only from confirmed outbreaks, which may result in underestimation of the number of cases
- Surveillance is likely not uniformly carried out in the different countries and there is likely variability in quality of surveillance within these countries.

METHODS

- The study was conducted in Kenya, Tanzania, Uganda, and Zambia
- District level confirmed cholera case data for several different time period were obtained for those countries
- District level population data of these countries were obtained from the respective government sources
- Population and case data were linked to district map of each of these countries
- Performed spatial analysis using SaTScan



- District/county level case and population counts with their geographic reference points (centroid of the district) were employed in the discrete Poisson model to identify significant clustering of cholera in the country
- Districts fell in the significantly clusters of higher risk for cholera were defined as the hotspots
- Populations living in these hotspots were then computed to determine the count of the at-risk population in these countries

POPULATION AT RISK

| Country | Total population | No. of total districts | Period studied | No. of cholera cases | No. of hotspot districts | Relative risk | No. of people in the hotspots | No. of people with relative risk >2.00 |
|--------------|--------------------|------------------------|----------------|----------------------|--------------------------|---------------|-------------------------------|--|
| Kenya | 47,848,478 | 47 | 2015-2018 | 24,711 | 13 | 1.24-13.43 | 13,942,722 | 3,128,238 |
| Uganda | 34,634,650 | 112 | 2011-2016 | 11,030 | 22 | 1.26-21.49 | 7,053,181 | 5,387,103 |
| Tanzania | 47,782,606 | 184 | 2015-2018 | 29,053 | 59 | 1.36-7.70 | 17,087,468 | 12,138,900 |
| Zambia | 13,092,666 | 72 | 2008-2017 | 34,950 | 16 | 1.50-78.87 | 4,690,121 | 3,057,121 |
| Total | 143,358,400 | 415 | -- | 99,744 | 110 | -- | 42,773,492 | 23,711,362 |

CONCLUSIONS

- The study identified cholera hotspots in the four countries in East and Southeast Africa
- Estimated the population at risk in the hotspot districts with various level of risk
- Magnitude of risk of having cholera in the hotspots could help guide how to move with the intervention strategy with the available resources
- Local knowledge is important for identifying the right places within the hotspot districts
- Targeting interventions at the right places followed by impact assessments of the interventions could eliminate cholera in that part of Africa

ACKNOWLEDGEMENT

Financial support for data analysis was provided by the Delivering Oral Vaccine Effectively (DOVE) project supported by the Bill and Melinda Gates Foundation (BMGF) and administered through the Johns Hopkins Bloomberg School of Public Health

Association between Flooding from Hurricane Katrina and Gentrification in Orleans Parish, Louisiana



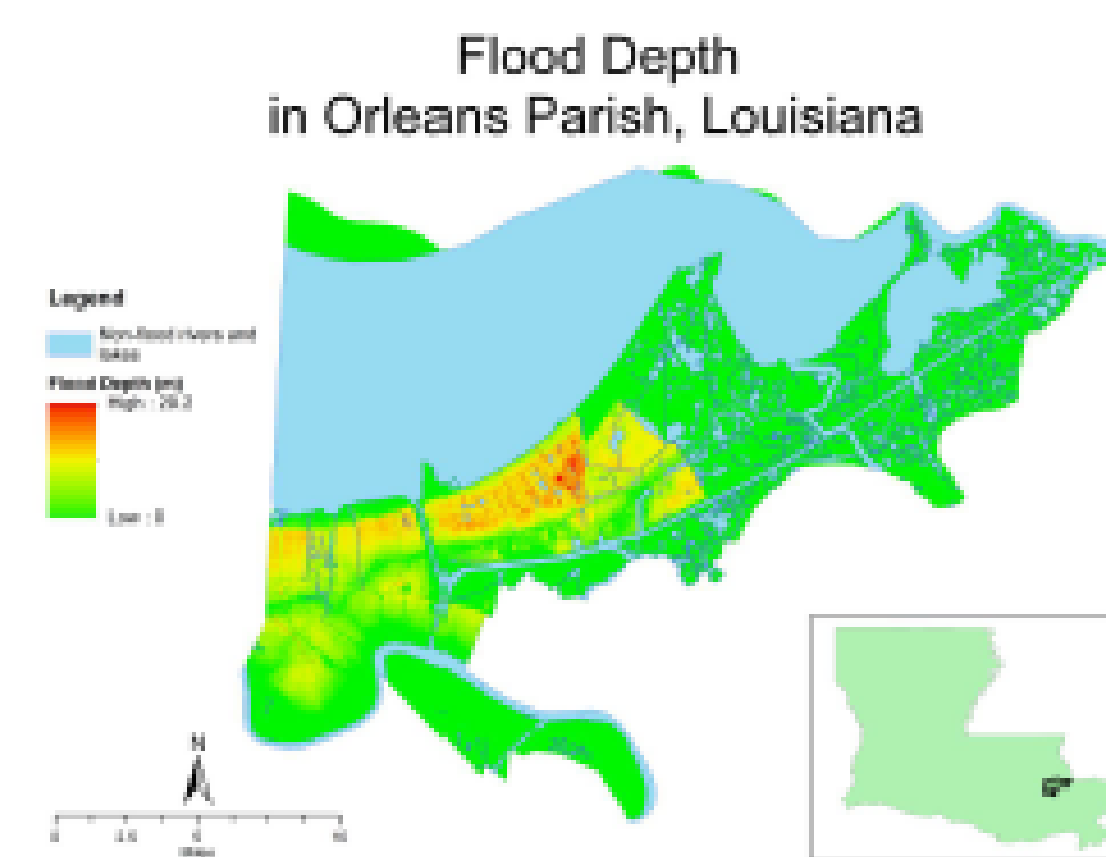
Kyle T. Aune, MPH,¹ Dean B. Gesch, PhD,² Genee S. Smith, PhD, MSPH¹
¹ Department of Environmental Health and Engineering, ² United States Geologic Survey

Background

- Hurricane Katrina, a Category 5 hurricane, affected the Gulf Coast in August 2005
- Much of the New Orleans area lies below sea level and the level of Lake Ponchartrain to the north
- Storm surge, levy failure, and the low-lying nature of the area led to widespread flooding and evacuation of 80-90% of the city's residents
- Rebuilding continues, but only 57% of the city's population of 175,000 Black residents returned
 - Some residents complain of gentrification, the change of a neighborhood from low value to high value that is associated with residential displacement and negative health effects
- This study investigates (1) the relationship between flooding severity and gentrification following Hurricane Katrina and (2) whether this relationship is modified by pre-storm Black race, the % of new homes constructed in an area, or the % of new residents in an area

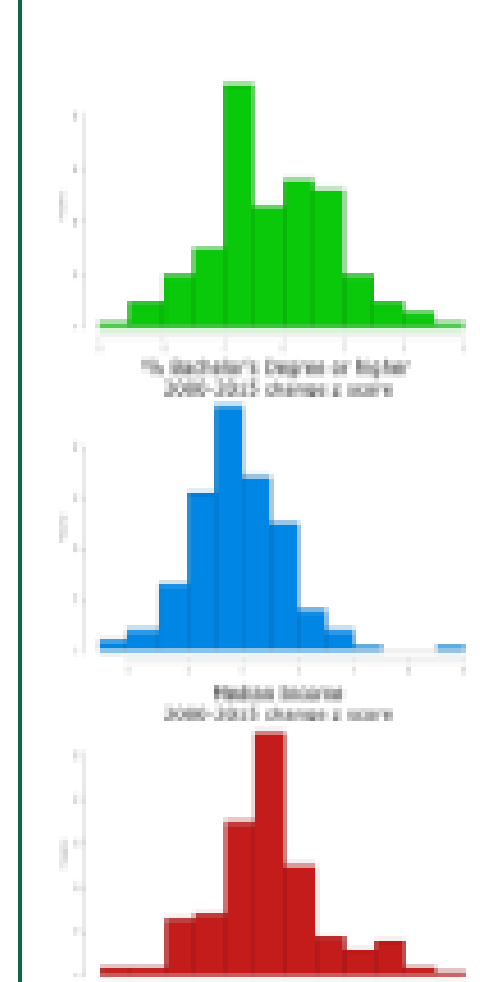
GIS Methods

- Light detection and ranging (LIDAR) data collected in 2002 prior to Hurricane Katrina to determine elevation
 - Original 5-meter resolution data processed into 10-m resolution for analysis
- Because of the unique layout of the New Orleans area, water gauge measurements at Lake Ponchartrain were able to be used to estimate the high water mark in the area during peak flooding after the hurricane
- Difference between elevation as measured by LIDAR and the gauge height calculated as flood depth
- Flood extent calculations visually confirmed against satellite imagery of peak flood extent
- Vector math used to calculate flood depth statistics by census tract
 - Because of underlying non-normal distribution of flood data, median depth calculated for use in analysis



Other Methods

- Demographic information about population race/ethnicity, education level, housing characteristics, and economic characteristics obtained from US Census records
 - 2000 decennial census, 2010 5-year American Community Survey (ACS) estimates (2006-2010), 2015 5-year ACS estimates (2011-2015)
- 2000 data cross-walked to reflect changes due to the 2010 re-drawing of census tract boundaries
- Three census tracts removed for zero population



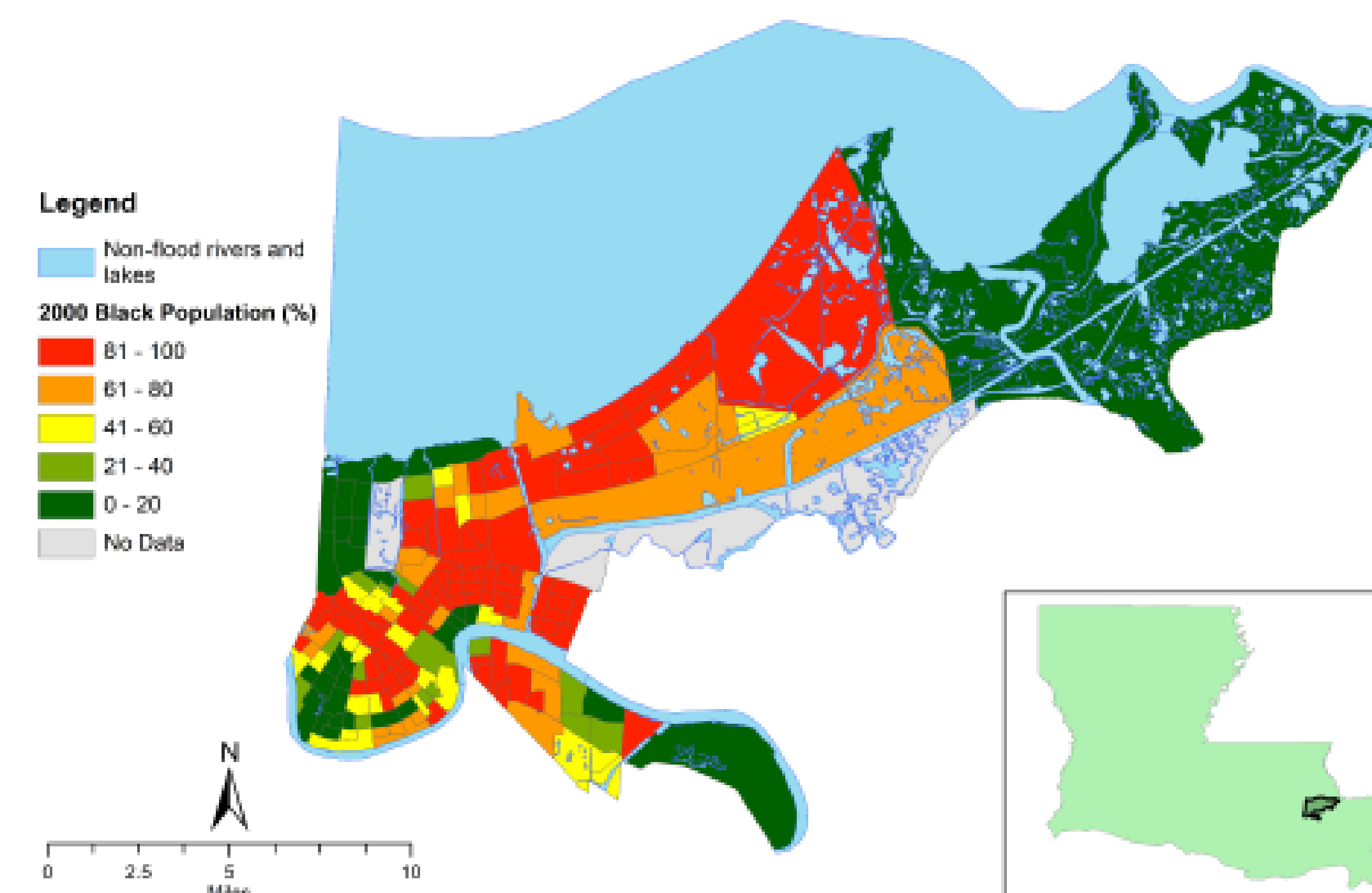
- Census tracts considered eligible to undergo gentrification if 2000 census tract median income was < 2000 parish median income (\$27,133)
- Gentrification index
 - Change from baseline to follow-up in % population with at least bachelor's degree, median income, % population above poverty limit
 - Z-score created for each change, summed to create overall gentrification index
- Binary variable created at ±0 to identify gentrification
- 4 categories (<-1 SD, -1 SD - 0, 0 - 1 SD, >1 SD)
- Binary logistic regression performed with product terms to test effect measure modification
- Base model confounders selected *a priori*
- Additional confounders identified through manual backward model selection using AIC optimization selection criteria

Population Demographics

| | 2000 | | p | 2006-2010 | | 2011-2015 | | p | |
|--------------------------------------|---------------------|------------------|---------|-------------------|-----------------------|-------------------|-----------------------|------------------|---------|
| | Not Eligible (N=79) | Eligible (N=95) | | Gentrified (N=44) | Not Gentrified (N=51) | Gentrified (N=53) | Not Gentrified (N=42) | | |
| Median flood depth (m) | 1.03 (1.05) | 0.70 (0.72) | 0.0194 | 0.44 (0.56) | 0.93 (0.76) | 0.0007 | 0.44 (0.61) | 0.91 (0.74) | 0.0009 |
| Population (N) | 3007.3 (1356.6) | 2601.0 (1130.9) | 0.0357 | 1400.5 (746.2) | 1478.0 (891.2) | 0.6458 | 1692.4 (673.8) | 1728.7 (1045.4) | 0.8380 |
| Race (%) | | | | | | | | | |
| Black | 44.0 (33.3) | 80.2 (22.5) | <0.0001 | 65.9 (31.2) | 78.4 (25.5) | 0.0373 | 51.5 (28.0) | 81.8 (19.5) | <0.0001 |
| White | 51.1 (33.3) | 15.6 (20.3) | <0.0001 | 30.1 (29.0) | 11.5 (14.1) | 0.0003 | 43.2 (25.9) | 10.5 (10.4) | <0.0001 |
| Other | 4.9 (3.5) | 4.2 (7.4) | 0.3921 | 4.0 (5.2) | 6.2 (14.0) | 0.3109 | 5.2 (4.0) | 5.8 (11.0) | 0.7470 |
| Hispanic (%) | 3.5 (1.6) | 3.0 (3.3) | 0.1855 | 5.5 (7.6) | 4.5 (7.6) | 0.5234 | 6.6 (5.2) | 5.7 (7.9) | 0.4976 |
| Education (%) | | | | | | | | | |
| High school diploma or higher | 86.1 (7.0) | 63.0 (12.5) | <0.0001 | 77.5 (14.0) | 72.2 (9.9) | 0.0415 | 83.3 (10.3) | 72.6 (9.9) | <0.0001 |
| Bachelor's degree or higher | 39.9 (17.9) | 14.4 (12.4) | <0.0001 | 26.4 (18.0) | 12.0 (11.5) | <0.0001 | 38.2 (18.4) | 12.4 (8.0) | <0.0001 |
| Residency (%) | | | | | | | | | |
| Homeowner | 50.5 (22.2) | 27.4 (14.1) | <0.0001 | 40.9 (19.5) | 37.0 (19.1) | 0.3341 | 36.2 (14.6) | 34.4 (18.1) | 0.5976 |
| Renter | 41.6 (20.0) | 65.8 (14.9) | <0.0001 | 53.5 (20.0) | 57.9 (20.5) | 0.2902 | 63.8 (14.6) | 65.6 (18.1) | 0.5976 |
| Dwelling (%) | | | | | | | | | |
| Single unit | 67.0 (21.6) | 55.3 (23.9) | 0.0009 | 53.0 (22.0) | 57.4 (25.4) | 0.3620 | 48.2 (17.1) | 57.9 (22.2) | 0.0190 |
| Multi unit | 29.4 (18.3) | 38.6 (20.7) | 0.0024 | 41.3 (18.5) | 36.2 (22.3) | 0.2195 | 46.4 (14.7) | 36.9 (19.0) | 0.0080 |
| Housing unit density (units/Ha) | 12.5 (7.5) | 19.8 (10.2) | <0.0001 | 14.3 (9.0) | 13.7 (10.0) | 0.7634 | 19.8 (9.6) | 12.8 (8.2) | 0.0003 |
| Moved in last 5 years (%) | | | | 52.6 (15.6) | 60.2 (16.1) | 0.0234 | 48.5 (10.0) | 43.2 (13.4) | 0.0345 |
| Home constructed in last 5 years (%) | | | | 4.3 (5.8) | 7.7 (19.6) | 0.2335 | 46.3 (9.9) | 41.9 (12.9) | 0.0664 |
| Median income (\$) | 43027.9 (19267.8) | 18152.2 (5311.6) | <0.0001 | 36781.8 (17701.8) | 19390.5 (6265.6) | <0.0001 | 25871.1 (9390.2) | 14939.3 (4860.6) | <0.0001 |
| Unemployment (%) | 5.7 (3.0) | 15.3 (9.4) | <0.0001 | 14.8 (11.2) | 19.0 (12.9) | 0.0948 | 10.7 (5.2) | 15.8 (7.8) | 0.0003 |
| Employment in management/sales (%) | 69.9 (12.6) | 48.0 (11.4) | <0.0001 | 49.2 (19.3) | 41.7 (18.4) | 0.0592 | 61.7 (12.3) | 43.1 (13.7) | <0.0001 |
| Population below poverty line (%) | 15.1 (7.4) | 40.7 (14.0) | <0.0001 | 24.5 (14.2) | 41.6 (13.9) | <0.0001 | 27.5 (11.9) | 45.1 (15.2) | <0.0001 |

Results

Black Race in Orleans Parish, Louisiana



Log odds of gentrification per 1 m increase in flood depth

| Model | OR (95% CI) |
|-----------------------------|------------------|
| Crude | 0.34 (0.17-0.68) |
| Base Model ^a | 0.45 (0.21-0.94) |
| Adjusted Model ^b | 0.40 (0.19-0.87) |

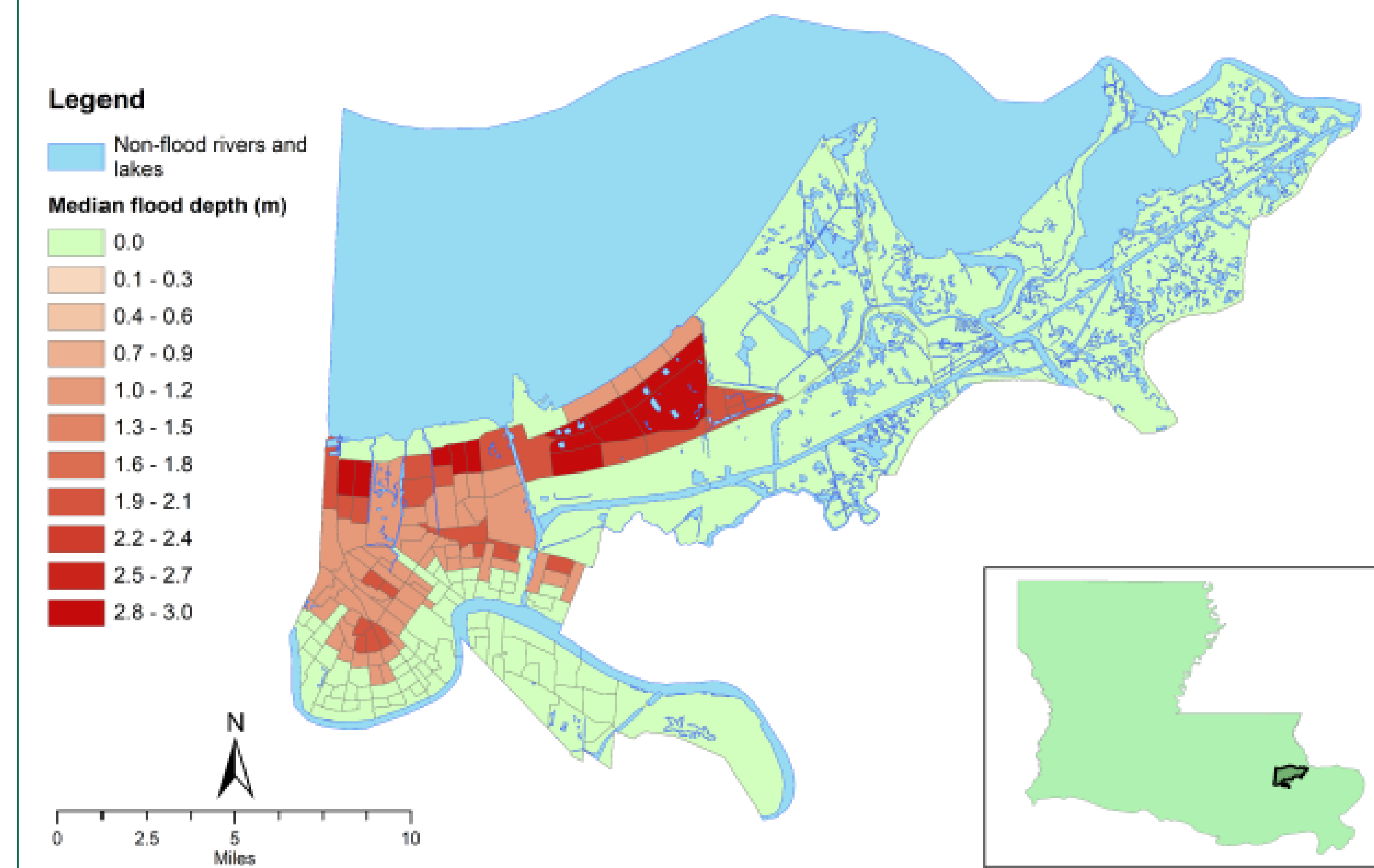
^a Adjusted for % Black population, % unemployment, % homeowners
^b in addition to base model adjustments, adjusted for % multi-unit dwellings, median home value

Stratified Analysis of Log Odds of Gentrification per 1 m increase in flood depth

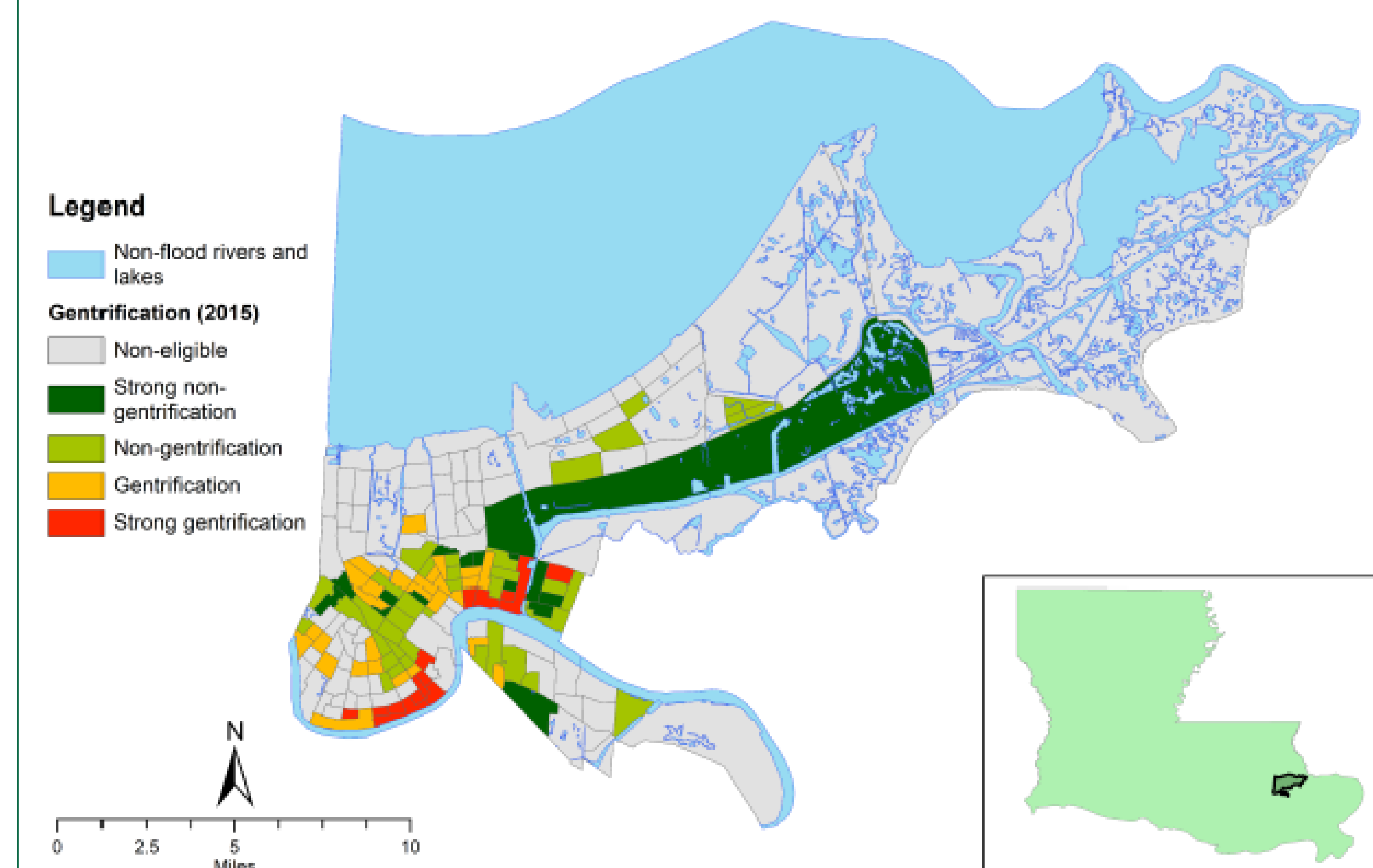
| Modifier | OR* (95% CI) | | Interaction P |
|--------------------------------------|------------------|------------------|---------------|
| | Bottom 50% | Top 50% | |
| Race | 0.25 (0.05-1.17) | 0.69 (0.21-2.24) | 0.0807 |
| Home constructed within last 5 years | 0.40 (0.12-1.36) | 0.43 (0.14-1.26) | 0.5127 |
| Moved in within last 5 years | 0.36 (0.10-1.25) | 0.50 (0.17-1.45) | 0.5309 |

* Adjusted for % Black population, % unemployment, % homeowners, % multi-unit dwellings, median home value

Flooding in Orleans Parish, Louisiana



Gentrification in Orleans Parish, Louisiana



Discussion & Future Directions

- Strong inverse relationship between gentrification and flood water depth, suggesting gentrification is less likely to occur in low-elevation areas that are flood prone
 - New property development likely to occur in areas of higher elevation
 - Keenan et al (2018) show that areas in Miami with low observed or expected flooding were more gentrified than flood-prone areas
- Future analysis will tease out the effects of elevation in addition to and in the absence of flooding after Katrina

Health care is local. So is health care transformation.

Utilizing ArcGIS to Improve Operational Decision Making in the Maryland Primary Care Program

Cameron Schilling^{1,2}, Leena Aurora^{1,2}

Johns Hopkins University Bloomberg School of Public Health¹; Maryland Department of Health, Program Management Office²



Introduction

The Maryland Primary Care Program (MDPCP) is an innovative health care transformation program developed by the Maryland Department of Health (MDH) and the Centers for Medicare and Medicaid Services (CMS).

The MDPCP provides funding and support for the delivery of advanced, comprehensive primary care. MDH's Program Management Office (PMO), in conjunction with the Center for Medicare and Medicaid Innovation (CMMI) are responsible for the MDPCP's implementation. ArcGIS was utilized to inform MDPCP decision-making within the PMO. Data was queried about MDPCP's primary care practices and innovative Care Transformation Organizations (CTOs). Operational decision making was improved by analysis and visualization of statewide spatial data. The interactive maps created by ArcGIS will serve as ongoing resources for program staff and leadership.

Objective

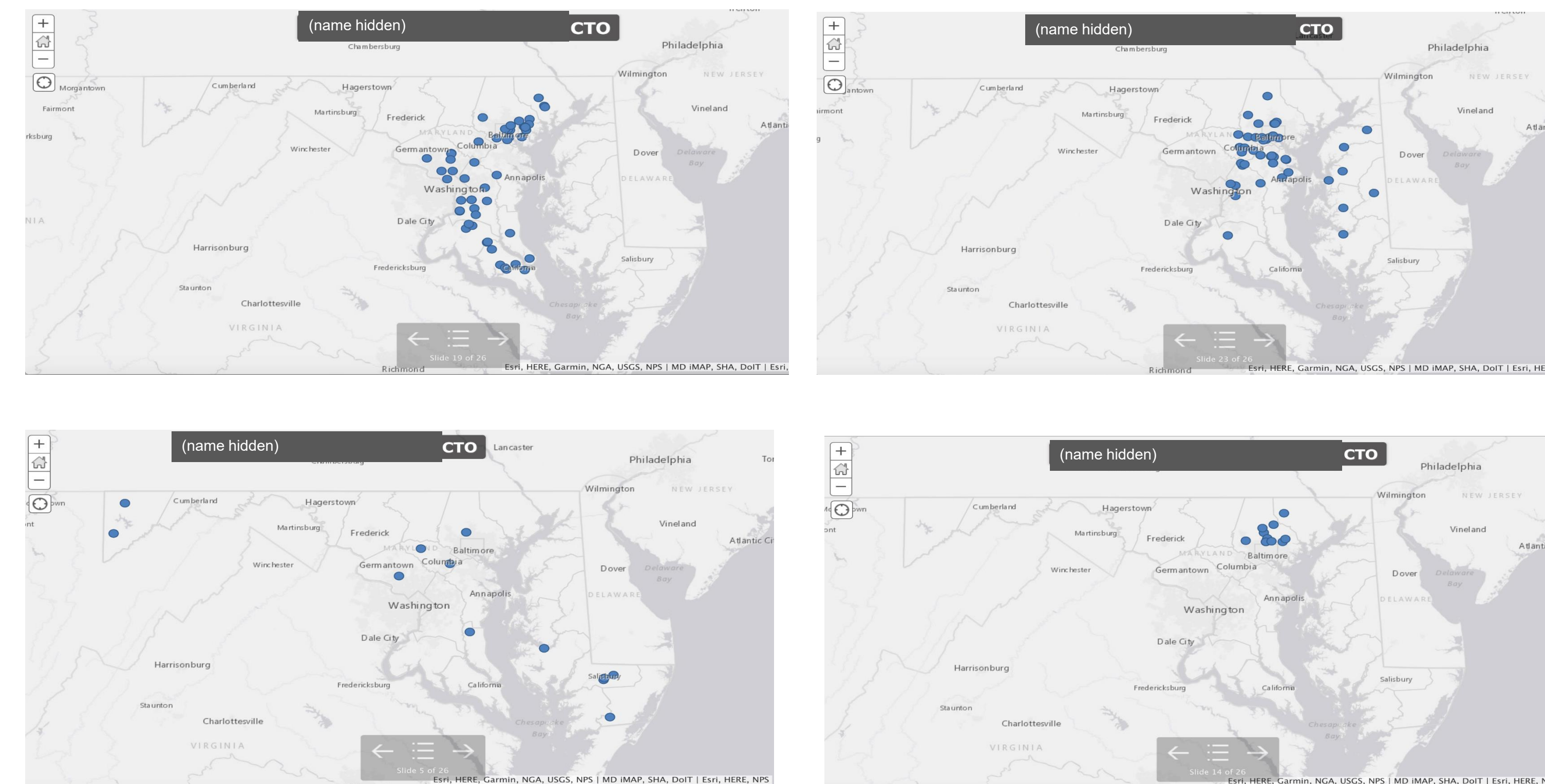
To use ArcGIS tools to optimize care transformation in Maryland via the MDPCP. In particular, the data that was visualized and manipulated regarded the allocation of primary care practices to their coaches, as well as the gaps and trends in the new Care Transformation Organizations (CTOs). These aimed to inform leadership in programmatic decision making and resource allocation.

Methodology

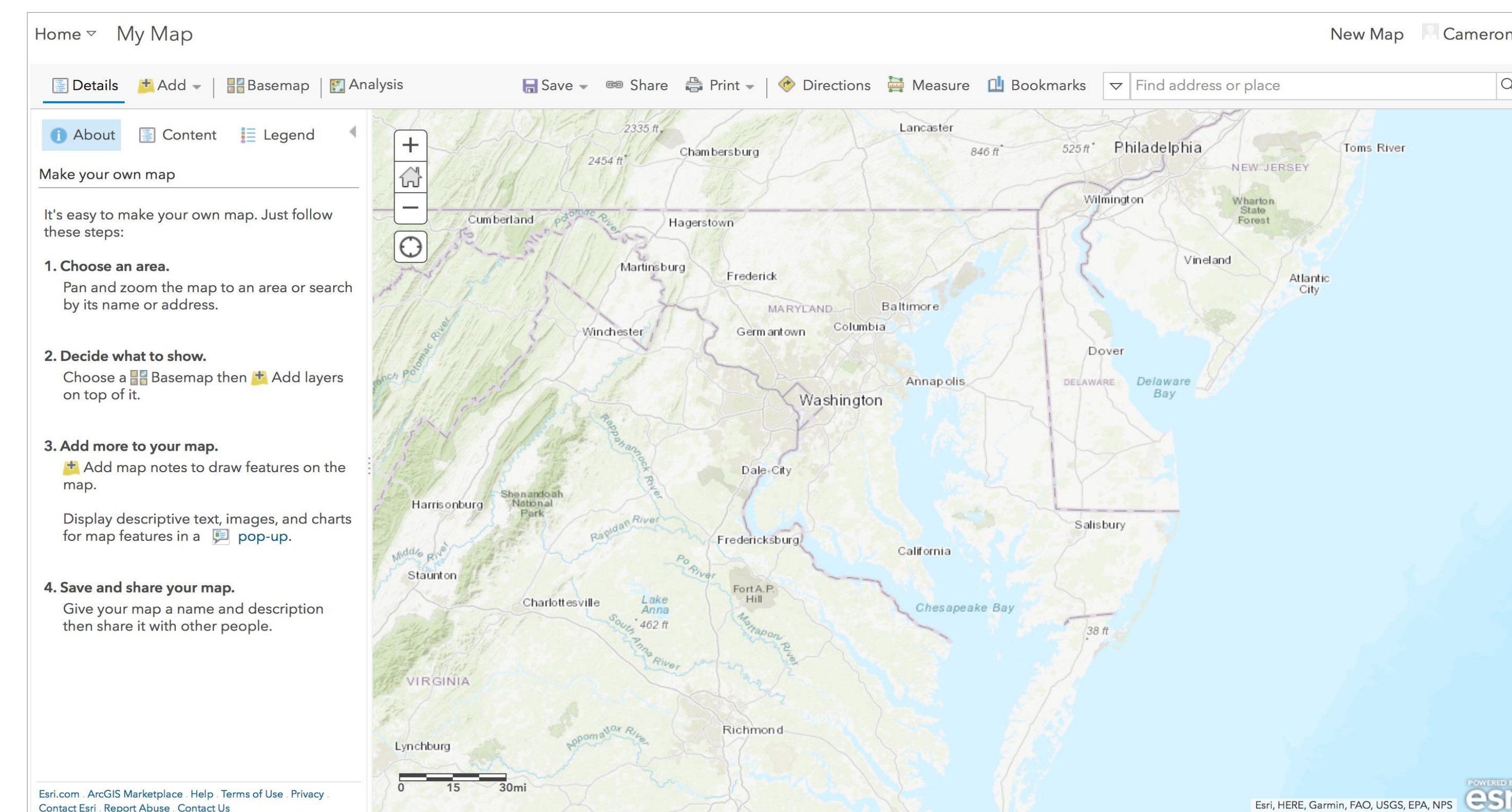
The Maryland Department of Health has developed an instance of the ArcGIS software that is made available to state employees. The ArcGIS instance allows access to a simplified graphical user interface and a robust set base map layers to utilize.

Upon program implementation, CMS and MDH collected data from primary care practices and CTOs via online applications for participation in MDPCP. Such data included information regarding location, preference for CTO, and track for which they wanted to apply. County data was not collected in the application and was obtained later through approximate county-zip mapping and manual address lookups. Practice applications included a list of providers that were part of that practice, allowing the PMO to conduct analyses based on practice size. Practice application data was aggregated with county-zip mappings and size of practice information. Ultimately, a single data table was generated that included relevant information for each practice. Dataset aggregation and cleansing was conducted with Microsoft Excel.

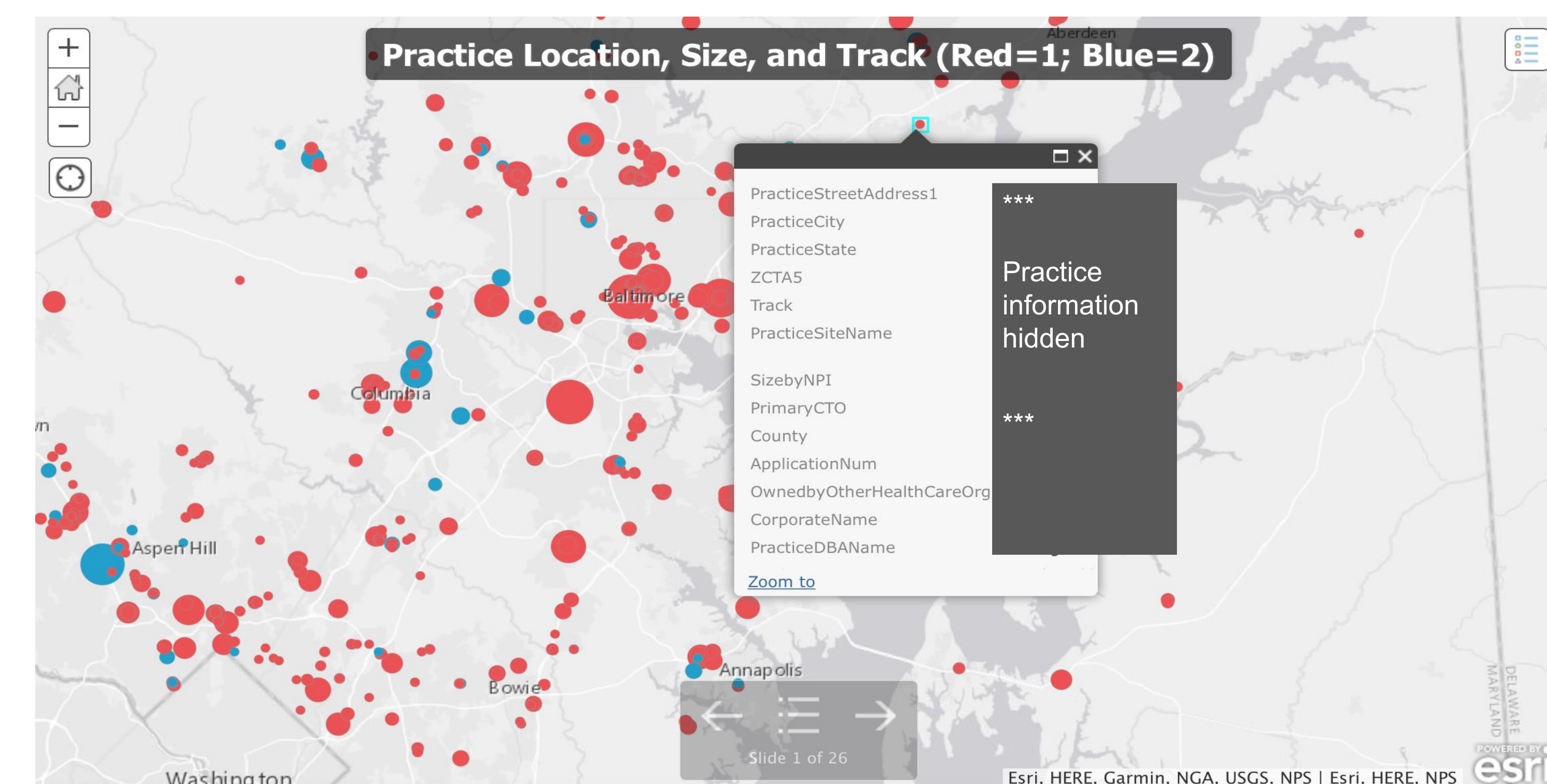
The master data table was loaded into the Maryland Department of Health ArcGIS web application as a map layer using addresses as geographic indicators. Analyses using Aggregate functions were conducted to generate multiple layers highlighting attributes of interest.



Comparing maps to view geographic spread and gaps of program CTOs.



Maryland state employees can access a web app that makes GIS intuitive and accessible; building in Maryland-specific data and geocoding tools.



Using the web apps' "presentation mode," program staff can easily view and compare detailed information about data points.

Results

ArcGIS maps reflecting practice and CTO locations are to be used as part of stakeholder engagement, resource allocation and forecasting, recruitment efforts, and relationship building.

Stakeholder Engagement: The operational analyses using ArcGIS were used in ongoing stakeholder engagement. The maps were a useful tool in communicating size, scope, trends, and potential risks within the Maryland Primary Care Program. Audiences and stakeholders included Care Transformation Organizations, the federal government, other state agencies, and state leadership. Maps are also planned to be shared with practices upon program implementation.

Resource Allocation: The enhanced understanding of practice location, size, and association with CTOs allowed for better prediction of geographic resource utilization and for the development of a framework for assigning practice coaches to individual practices. Counties with high density practices and larger practices are expected to require more attention from the state as it implements the program. Practices that did not select a CTO are also predicted to require more attention from the state.

An initial framework for assigning practice coaches to practices evenly distributed the non-CTO practices into regions with consistent number of practices and a similar number of providers. The preliminary assignments were done by hand and in excel with the aid of maps showing practice locations, number of practices per county, and number of providers per county, limited to those practices without a CTO. A revised framework is under development to include practices with CTOs and home addresses of practice coaches. The same considerations will be taken into account with aims at maximizing efficiency and consistency between coaches for the amount of time available to assist practices.

Recruitment Efforts: Approximately 80% of the primary care practices within the state have applied for participation in this program for 2019. In future years, the state would like to recruit additional primary care practices to participate. Maps are being used to identify where gaps exist in program participation to indicate where recruitment efforts should be targeted.

Relationship Building: The ArcGIS set of maps used for operational decision making are also to be used to familiarize practice coaches with practice attributes. Many practices will be assigned to a single practice coach. Maps specific to each coach will be developed as an aid to build understanding of who the practices are that they will be visiting on site. Practices that are clustered together are predicted to have similar patient populations and may be targeted to engage as a more tailored learning network within MDPCP.

Conclusion

The MDPCP supports the overall health care transformation process and allows primary care providers to play an increased role in prevention, management of chronic disease, and preventing unnecessary hospital utilization. Because the program is in its first year, many important decisions are being made and many factors need to be considered to best serve the state.

These maps are useful for informing policy makers and others about geographic scope and gaps in program resources. As the program is implemented, these will continued to be utilized within the program to inform resource utilization, coach assignments and future recruitment efforts. They will also be used as "on the ground" data visualization tools for stakeholder engagement and practice coaching.

Acknowledgements

We would like to thank the Maryland Department of Health PMO team for their support and mentorship. We especially thank program leadership: Dr. Howard Haft, Chad Perman, Alice Sowinski-Rice, and Rachel Maxwell.

The Changing Epidemiology of Non-Fatal Overdose in Baltimore City: Insights from Emergency Medical Services Encounters

Chen Dun, Brian W. Weir
 Department of Health, Behavior and Society
 Johns Hopkins Bloomberg School of Public Health

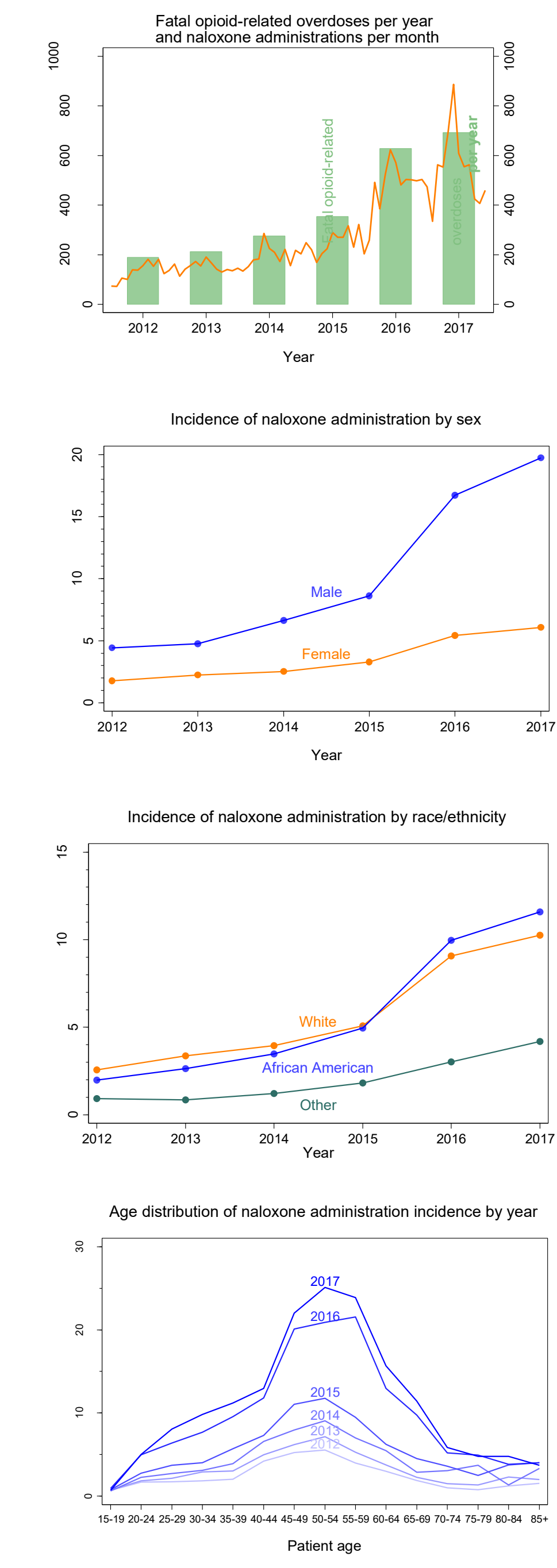
Background

- Opioid-related overdose:**
- Fatal overdose increased five-fold from 1999 to 2016 in the United States, reaching 72,885 in 2017 (CDC, 2018).
 - In Baltimore City, fatal opioid-related overdose increased from 354 in 2015 to 628 in 2016 (Maryland Dept. of Health, 2018).
- Baltimore City EMS:**
- Baltimore City Fire Department (BCFD) operates the city's single EMS system
 - EMS medics are often first responder to overdoses
 - EMS electronic records provide demographic, temporal and spatial information
- Emergency medical services (EMS) collect data that:**
- Provides real-time, incidence-based surveillance of violence
 - Provides population-based incidence rates
 - Reveals individual-level correlates
 - Reveals geospatial patterns and correlates

Methods

- EMS data collection and coding:**
- Baltimore City Fire Department EMS records (Jan 2012 -- Dec 2017)
 - Naloxone administration for age older than 15 (N = 20,592)
- Descriptive analyses:**
- City-level population-based incidence rates per 1,000 persons older than 15
 - Univariable and multivariable negative binomial regression
- Geospatial analyses:**
- Spatial join to Baltimore City census tracts
 - Mapping of census tract incidence rates and approximate incidence locations in 2013, 2015 and 2017

Results



- Overall:**
- EMS encounters with naloxone administration followed similar pattern to fatal opioid-related overdoses in Baltimore (Maryland Dept. of Health, 2018). Both increased over time with greater increase after 2015.
 - Naloxone administrations approximately 12-fold higher than fatal opioid-related overdoses.
- Sex:**
- The incidence rates for naloxone administration increased for both males and females.
 - Incidence rates for males approximately twice as high as for females.
 - Over time (see table) incident rate ratios were higher for males than females.
- Race:**
- Similar incidence rates for African Americans and Whites. Incidence rate increased significantly more for African Americans vs. Whites over time (see table).
 - Incidence rate lowest for "other" race/ethnicity. This may be due to different categories and differential misclassification for medic-recorded EMS encounters vs. self-reported census.
- Age groups:**
- Incidence rates were highest for middle-age (45-59) over all calendar years.
 - In 2013, only 50-54 age group had incidence over 7. By 2017, all age groups from 25 to 69 had incidence over 7 (see table).

- Characteristics of EMS encounters:**
- Mean records of naloxone administration per encounter = 1.13 (SD = .38; range = 1 – 8)
 - Mean total dose per encounter = 1.85 mg (SD = .66)
 - After naloxone administration, patient condition improved for 79%

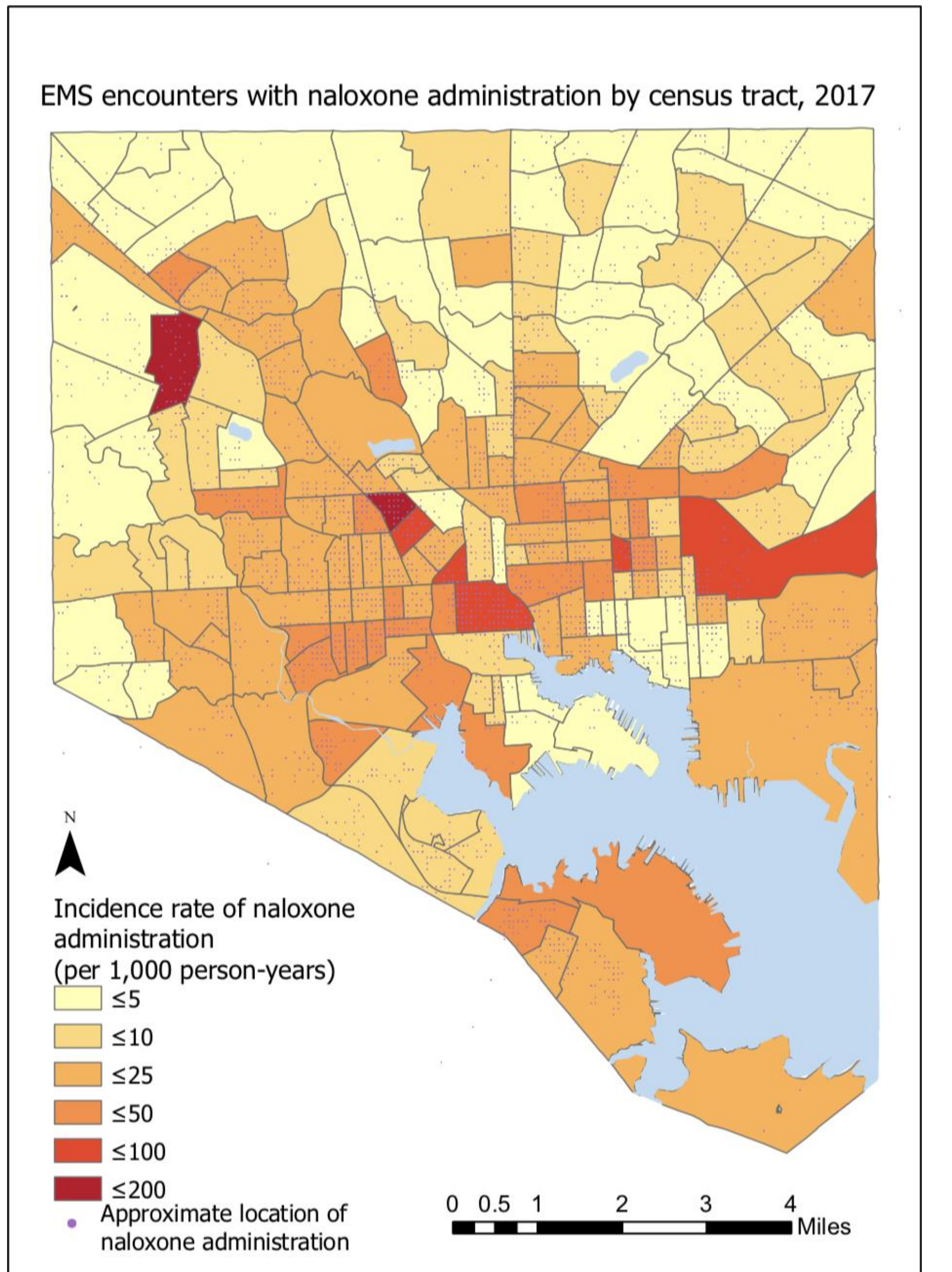
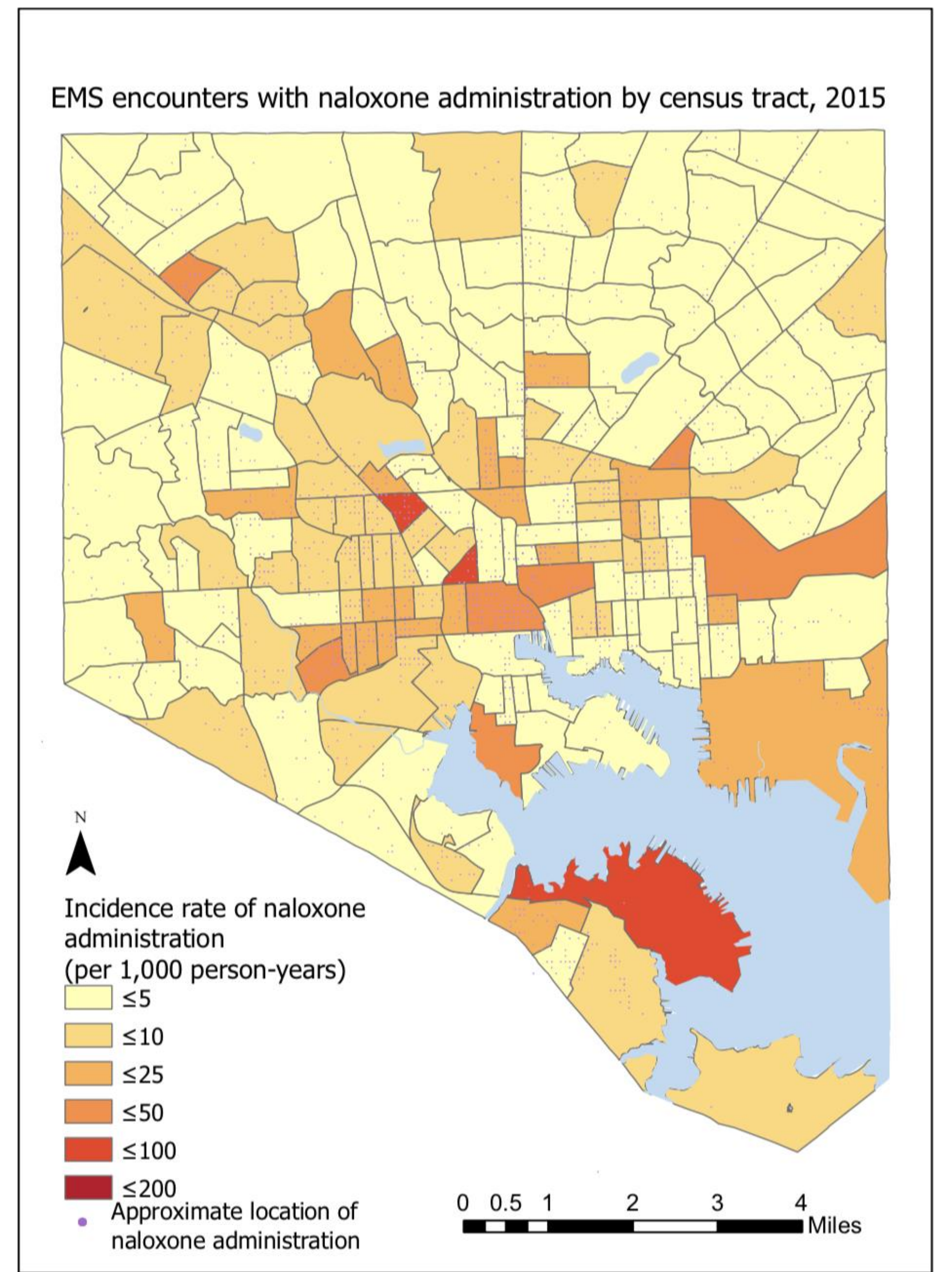
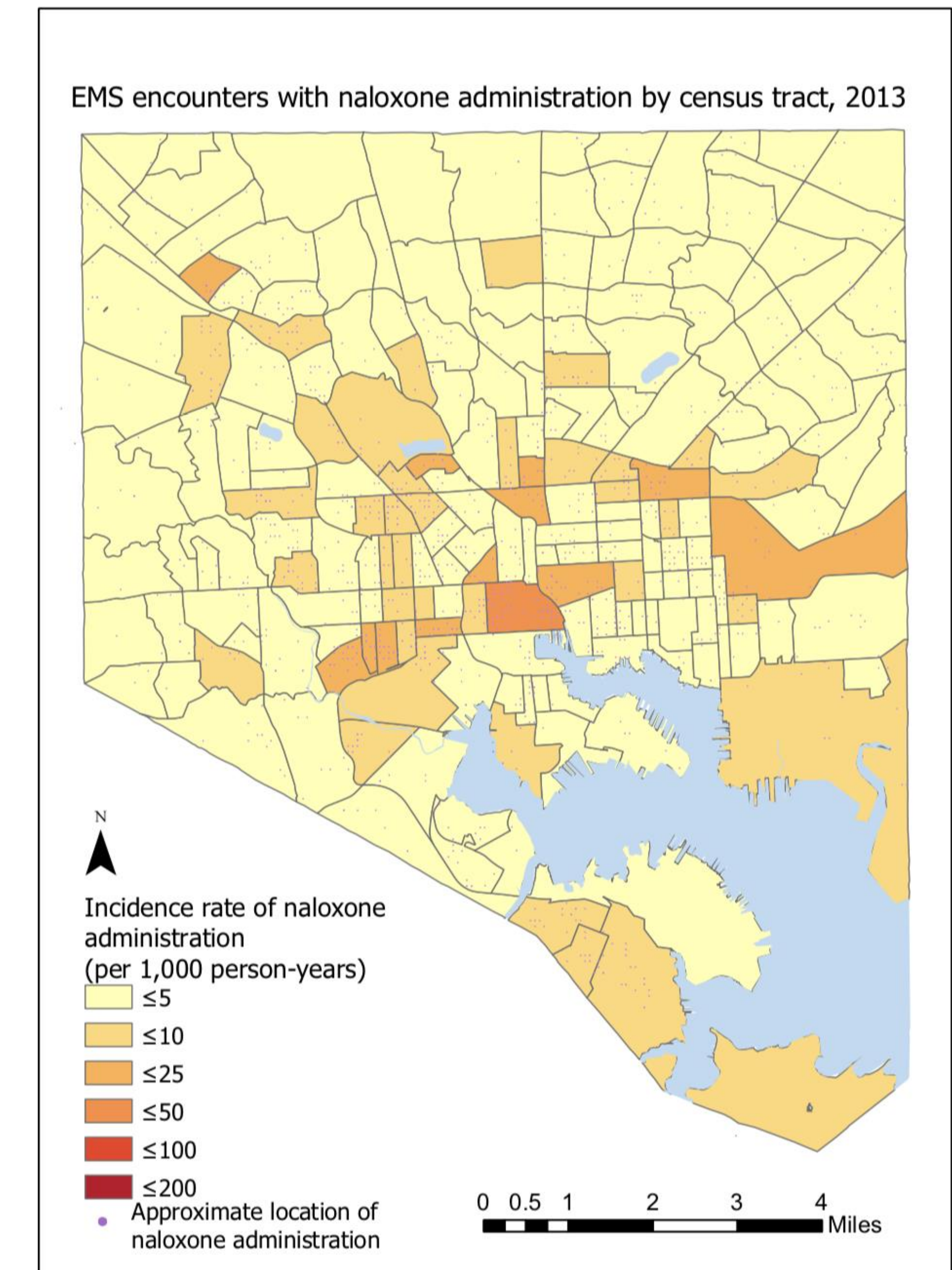
Incidence rates (IR) of naloxone administration per 1,000 person-years and incidence rate ratios (IRR) for 2015 vs. 2013 and 2017 vs. 2015.

| | 2013 | | 2015 | | 2017 | | 2015 vs. 2013 | | 2017 vs. 2015 | |
|-----------------------|------|------------|------|--------------|------|--------------|---------------|------------|---------------|------------|
| | IR | (95% CI) | IR | (95% CI) | IR | (95% CI) | IRR | (95% CI) | IRR | (95% CI) |
| Sex | | | | | | | | | | |
| Female | 2.2 | (2.1, 2.4) | 3.3 | (3.1, 3.5) | 6.1 | (5.8, 6.4) | 1.5 | (1.3, 1.6) | 1.8 | (1.7, 2.0) |
| Male | 4.8 | (4.5, 5.0) | 8.6 | (8.3, 9.0) | 19.7 | (19.2, 20.3) | 1.8 | (1.7, 2.0) | 2.3 | (2.2, 2.4) |
| Race/ethnicity | | | | | | | | | | |
| White | 3.4 | (3.1, 3.7) | 5.1 | (4.7, 5.4) | 10.3 | (9.8, 10.8) | 1.5 | (1.3, 1.7) | 2.0 | (1.9, 2.2) |
| African American | 2.6 | (2.5, 2.8) | 5.0 | (4.7, 5.2) | 11.6 | (11.2, 12.0) | 1.9 | (1.7, 2.0) | 2.3 | (2.2, 2.5) |
| Other | 0.9 | (0.6, 1.2) | 1.8 | (1.5, 2.3) | 4.2 | (3.6, 4.8) | 2.1 | (1.4, 3.1) | 2.3 | (1.8, 3.0) |
| Age | | | | | | | | | | |
| 15-19 | 0.7 | (0.5, 1.1) | 0.8 | (0.5, 1.1) | 0.7 | (0.5, 1.0) | 1.0 | (0.6, 1.8) | 0.9 | (0.5, 1.6) |
| 20-24 | 1.9 | (1.5, 2.3) | 2.8 | (2.3, 3.3) | 5.0 | (4.4, 5.7) | 1.5 | (1.1, 2.0) | 1.8 | (1.5, 2.2) |
| 25-29 | 2.2 | (1.8, 2.6) | 3.7 | (3.3, 4.2) | 8.1 | (7.4, 8.8) | 1.7 | (1.4, 2.1) | 2.2 | (1.9, 2.5) |
| 30-34 | 2.9 | (2.5, 3.4) | 4.0 | (3.5, 4.6) | 9.8 | (9.0, 10.7) | 1.4 | (1.1, 1.7) | 2.4 | (2.1, 2.9) |
| 35-39 | 3.0 | (2.5, 3.6) | 5.7 | (5.0, 6.5) | 11.2 | (10.2, 12.3) | 1.9 | (1.5, 2.4) | 2.0 | (1.7, 2.3) |
| 40-44 | 5.0 | (4.3, 5.8) | 7.3 | (6.5, 8.3) | 13.0 | (11.8, 14.2) | 1.5 | (1.2, 1.8) | 1.8 | (1.5, 2.1) |
| 45-49 | 6.2 | (5.4, 7.0) | 11.0 | (10.0, 12.2) | 22.0 | (20.5, 23.6) | 1.8 | (1.5, 2.1) | 2.0 | (1.8, 2.2) |
| 50-54 | 7.1 | (6.4, 8.0) | 11.8 | (10.8, 12.9) | 25.1 | (23.6, 26.7) | 1.6 | (1.4, 1.9) | 2.1 | (1.9, 2.4) |
| 55-59 | 5.3 | (4.6, 6.0) | 9.5 | (8.6, 10.5) | 23.9 | (22.4, 25.4) | 1.8 | (1.5, 2.1) | 2.5 | (2.2, 2.8) |
| 60-64 | 3.7 | (3.2, 4.4) | 6.2 | (5.5, 7.1) | 15.7 | (14.4, 17.0) | 1.7 | (1.3, 2.1) | 2.5 | (2.2, 2.9) |
| 65-69 | 2.2 | (1.8, 2.9) | 4.5 | (3.8, 5.4) | 11.4 | (10.2, 12.8) | 2.0 | (1.5, 2.7) | 2.5 | (2.0, 3.1) |
| 70-74 | 1.5 | (1.0, 2.2) | 3.6 | (2.8, 4.6) | 5.9 | (4.9, 7.1) | 2.4 | (1.5, 3.7) | 1.6 | (1.2, 2.2) |
| 75-79 | 1.4 | (0.9, 2.2) | 2.5 | (1.8, 3.5) | 4.8 | (3.7, 6.1) | 1.8 | (1.0, 3.3) | 1.9 | (1.3, 2.9) |
| 80-84 | 2.3 | (1.5, 3.5) | 3.7 | (2.7, 5.2) | 4.8 | (3.6, 6.4) | 1.6 | (1.0, 2.8) | 1.3 | (0.8, 2.0) |
| 85 or older | 2.0 | (1.3, 3.1) | 4.0 | (3.0, 5.4) | 3.7 | (2.7, 5.1) | 2.0 | (1.2, 3.4) | 0.9 | (0.6, 1.4) |

- Geographic distribution:**
- Incidence increased dramatically in many census tracts. Centrally-located census tracts had highest incidence in 2017. Incidence was lower among census tracts near the western, northern, and eastern city borders.

Conclusions

- Summary:**
- Incidence of naloxone administration increased dramatically over time, and mirrored increases in fatal opioid-related overdoses.
 - Incidence increased for all groups defined by sex, race/ethnicity, and age.
 - Incident rate ratios over time were higher for males and non-Whites.
 - Absolute increase in incidence was higher among demographic groups with already elevated incidence.
- Limitations:**
- In some cases, naloxone administration may not indicate an opioid-related overdose, although high rates of patient improvement (79%) suggests exposure.
 - EMS records may be prone to misclassification error.
 - Incidence rate in city center may be inflated due to presence of non-residents.
- Recommendations:**
- EMS records may be used as a proxy for fatal opioid-related overdose trends.
 - Neighborhood factors associated with overdose should be evaluated in multivariable models.
 - EMS encounters with naloxone administration may be suitable for naloxone distribution and harm reduction interventions.
- Acknowledgements:**
- Baltimore City Fire Department Emergency Medical Services (Kelly King, Chief James Matz, Chief Mark Fletcher). ArcGIS Pro help from Bonnie Wittstadt.
 - Johns Hopkins University Center for AIDS Research (P30AI094189)
 - Contacts: Chen Dun (cdun1@jhmi.edu), Brian W. Weir (bwreir3@jhu.edu)



Objective

Poor access to local urology care has been associated with high urologic cancer mortality rates (MR). Telemedicine, which relies on the use of high-speed internet to provide remote healthcare services to patients, has been proposed as a strategy to decrease geographic disparities in urologic cancer MR.

However, the current relationship between internet access and urologic outcomes is not well understood. In this study we aim to establish how broadband internet access is associated with urologic cancer mortality rates (MR) for all US counties

Further, this work identifies counties with high urologic cancer mortality rates that could most benefit from telemedicine today: where there are already sufficient internet capabilities, and in the future, with the expansion of broadband internet services.

Materials and Methods

Univariate (unadjusted) and multivariate (adjusted) regressions were run to assess the association between 2014 county-level MR for bladder (BC), kidney (KC), prostate (PC), and testis cancer (TC) and county-level internet access rates (IAR). MRs were calculated via a validated Bayesian small-area estimation model with death records from the National Center for Health Statistics.

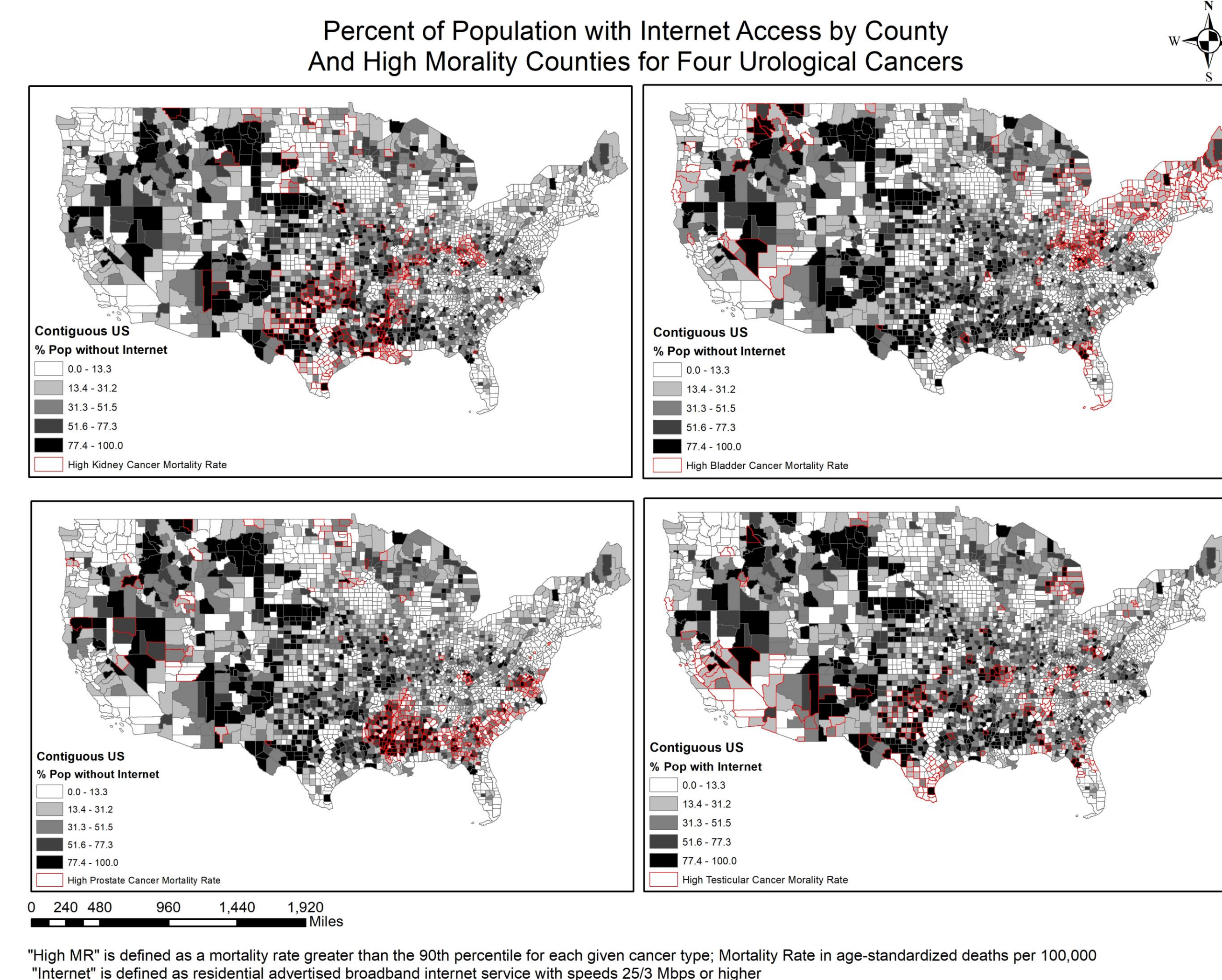
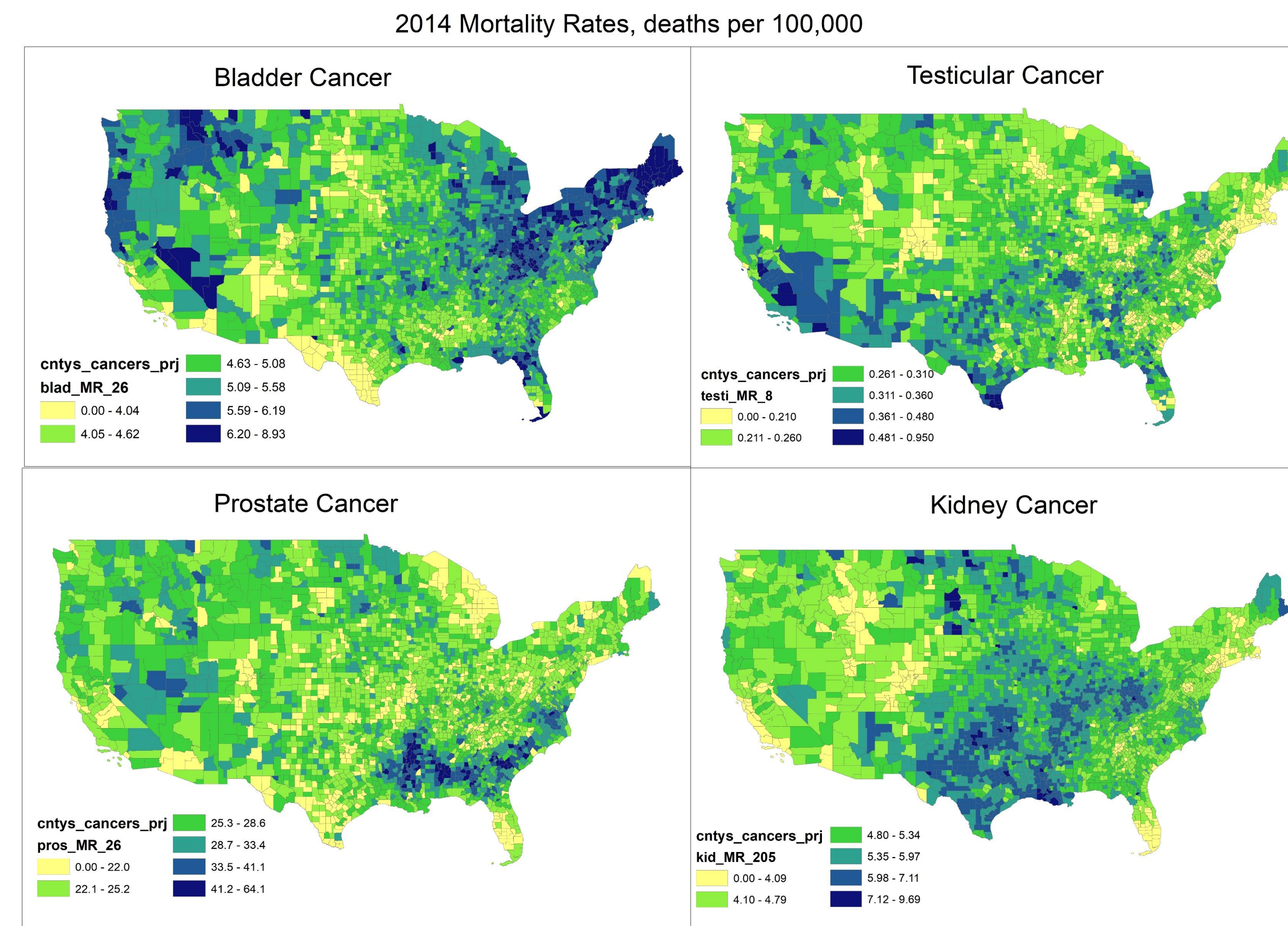
Counties with a cancer MR greater than the 90th percentile for a particular urologic cancer were targeted for potential telemedicine intervention. We also highlight counties that have poor access to local urologic care, defined as having zero urologists in a county, and, we examined the percentage of the population with access to broadband internet as of 2016. Finally, state-level data on telemedicine regulations were incorporated to specify counties best suited for telemedicine interventions based on existing legislative policies

Table 1. Results of Selected Regression Models for County-Level Analysis. The reported significant effects are the estimated changes in Bayesian smoothed age-adjusted urologic cancer mortality rates associated with each variable on the left. Positive coefficients indicate increase in mortality rate, more deaths per 100,000 people, per unit change of the variable indicated (n = 3140).

| Variable | Bladder Cancer | | Kidney Cancer | | Prostate Cancer | | Testis Cancer | |
|--|----------------|---------|---------------|---------|-----------------|---------|---------------|---------|
| | Beta | p-value | Beta | p-value | Beta | p-value | Beta | p-value |
| (Intercept) | 3.88 | <0.001 | 1.68 | <0.001 | 32.6 | <0.001 | 0.577 | <0.001 |
| % Population without internet ¹ | -0.015 | <0.001 | 0.019 | <0.001 | 0.034 | 0.23 | 0.001 | 0.015 |
| Number of urologists | 0.002 | 0.01 | -0.003 | <0.001 | -0.026 | <0.001 | -- | -- |
| Urban-rural continuum ² | -0.059 | <0.001 | 0.014 | 0.006 | -- | -- | -- | -- |
| % Population non-Hispanic white ¹ | 0.16 | <0.001 | 0.042 | <0.001 | -0.942 | <0.001 | 0.006 | <0.001 |
| Median household income (\$1,000 USD) ³ | 0 | <0.001 | 0 | <0.001 | 0 | <0.001 | -0.001 | <0.001 |
| % Population high school grad or higher ¹ | -- | -- | -- | -- | 0.794 | <0.001 | -0.032 | <0.001 |
| COPD ⁴ mortality rate, 2014 | 0.012 | <0.001 | -- | -- | -- | -- | -- | -- |
| Obesity prevalence (% of population) | -- | -- | 0.086 | <0.001 | -- | -- | -- | -- |
| HHD ⁵ mortality rate, 2014 | -- | -- | 0.004 | 0.011 | -- | -- | -- | -- |

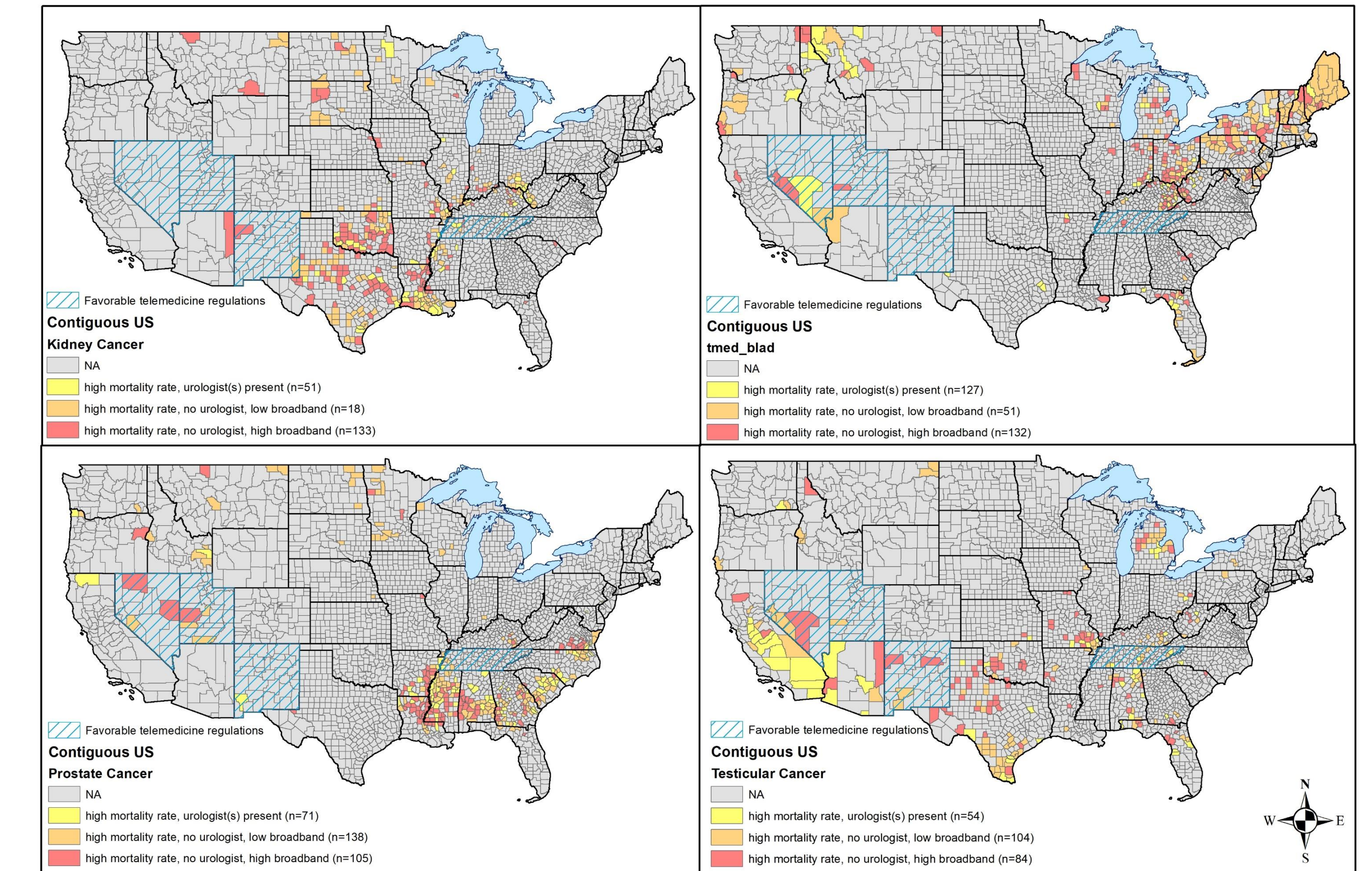
1. Unit changes in percent of population are in 10% increments.
2. Rural urban index, established by the US Department of Agriculture, is a classification scheme that distinguishes metropolitan counties by the population size of their metro area, and nonmetropolitan counties by degree of urbanization and adjacency to a metro area. The value ranges from 1 most urban to 9 most rural, and it is treated here as a continuous variable.
3. Effects shown of 0.000 are actually <0.001.
4. Bayesian smoothed mortality rates for COPD, chronic obstructive pulmonary disorder, are used as a proxy for smoking status at the county level.

Results



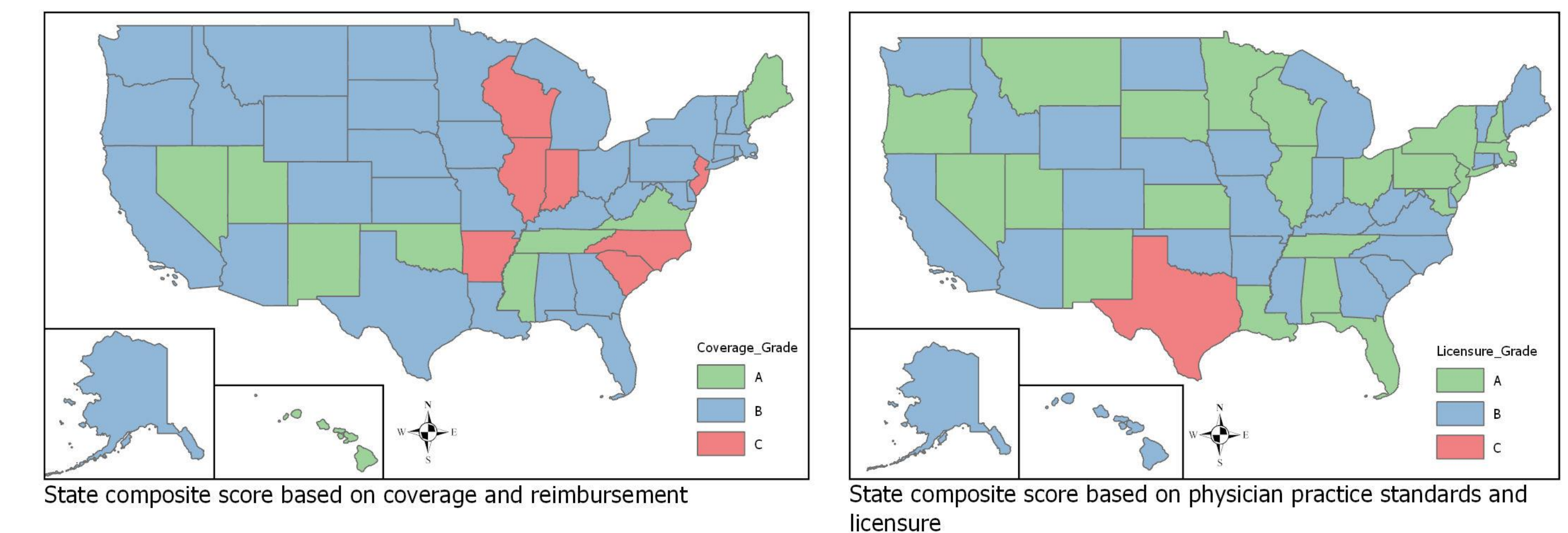
High MR is defined as a mortality rate greater than the 90th percentile for each given cancer type; Mortality Rate in age-standardized deaths per 100,000
Internet is defined as residential advertised broadband internet service with speeds 25/3 Mbps or higher

Telemedicine Intervention Areas High Mortality for Four Urologic Cancers



High mortality rate is defined as a mortality rate greater than the 90th percentile for each given cancer type
High broadband is defined as 50% or more of the population in a county having access to residential advertised broadband internet service with speeds 25/3 Mbps or higher
Favorable telemedicine regulations identify states with an A rating in assessments of coverage practices and physician licensure requirements by the American Telemedicine Association

State Telemedicine Gaps Analysis Composite Scores



Conclusions

An increase in broadband internet access is associated with an increase in county-level kidney cancer mortality, but with a decrease in bladder cancer mortality.

397 total counties have high MR in at least one urologic cancer, no local access to a local urologist, and adequate internet access. 23 of these counties are located in states with telemedicine regulation "A" ratings. 99 counties had adequate access to broadband internet, no access to a local urologist, and had high mortality rates in at least two of four urologic cancers. 5 of those counties are in states with telemedicine regulation "A" ratings.

Urologists interested in using telemedicine to address geographic disparities in cancer mortality should consider targeting the counties identified in our analysis for future interventions.

Can multispectral information improve remotely sensed estimates of total suspended solids? A statistical study in Chesapeake Bay

Nicole M. DeLuca^{1*}, Benjamin F. Zaitchik¹, Frank C. Curriero² ¹Johns Hopkins University, Department of Earth and Planetary Sciences, Baltimore, MD

*ndeluca1@jhu.edu

²Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology, Baltimore, MD

Introduction

- Particles in water column limit light availability for Bay grasses and benthic organisms
- USEPA Total Maximum Daily Load (TMDL) targets for Chesapeake Bay include total suspended solids (TSS)
- Monitoring TSS through *in situ* sampling can be limited by time and funding – remote sensing improves spatial and temporal coverage
- Many MODIS algorithms for TSS utilize red and NIR wavelengths shown to correlate well with TSS concentration (red bars in Figure 1)
- **Motivation:** Using eight statistical and machine learning models, we investigate whether remotely sensed TSS estimates in the Chesapeake Bay can be improved by utilizing additional MODIS wavelengths (gray bars in Figure 1).

Table 1. MODIS ocean color products used in this study and shown as bars in Figure 1.

| | |
|-----------------------------|-----------------------------|
| Rrs_412 (sr ⁻¹) | Rrs_555 (sr ⁻¹) |
| Rrs_443 (sr ⁻¹) | Rrs_645 (sr ⁻¹) |
| Rrs_469 (sr ⁻¹) | Rrs_667 (sr ⁻¹) |
| Rrs_488 (sr ⁻¹) | Rrs_678 (sr ⁻¹) |
| Rrs_531 (sr ⁻¹) | Rrs_859 (sr ⁻¹) |
| Rrs_547 (sr ⁻¹) | - |

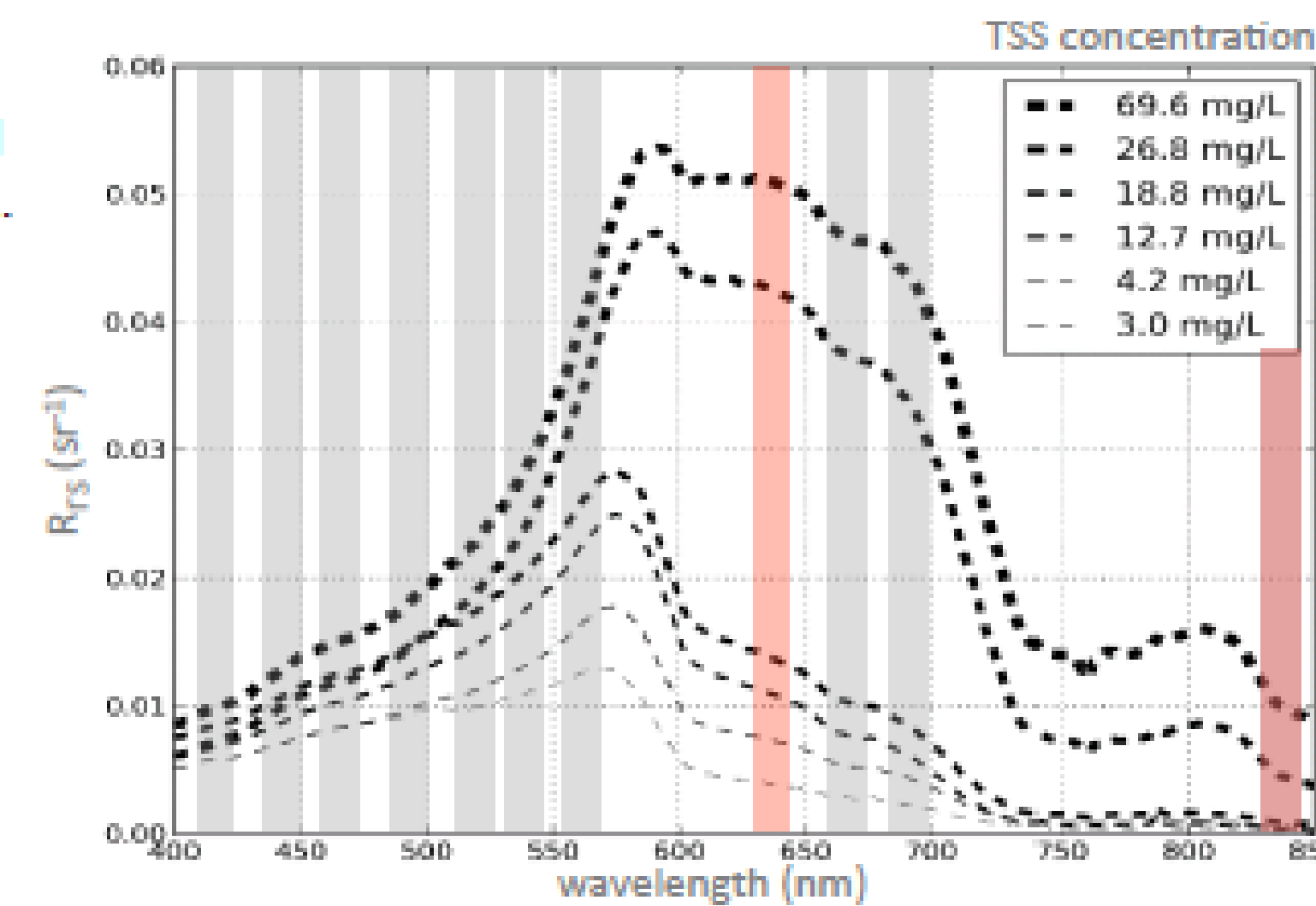


Figure 1. Spectra of TSS concentrations in coastal waters. Adapted from Dorji et al., 2016²

Methods

Satellite data processing

- MODIS Aqua images from 2003-2016
- Batch processed to Level 2 in SeaDAS 7.4
- Standard iterative NIR atmospheric correction
- All products processed at 1 km resolution (Table 1)

In situ measurements

- In situ data from Chesapeake Bay Program (CBP)
- Included measurements taken at or above top 1 meter of water column

Satellite-in situ matchups

- MODIS pixel and CBP station matchups on same day and within 250 meters
- Excluded matchups if one or more of the 11 MODIS bands (Table 1) were saturated
- Matchups (n=1360) cover 490 days, 86 station locations (Figure 2)

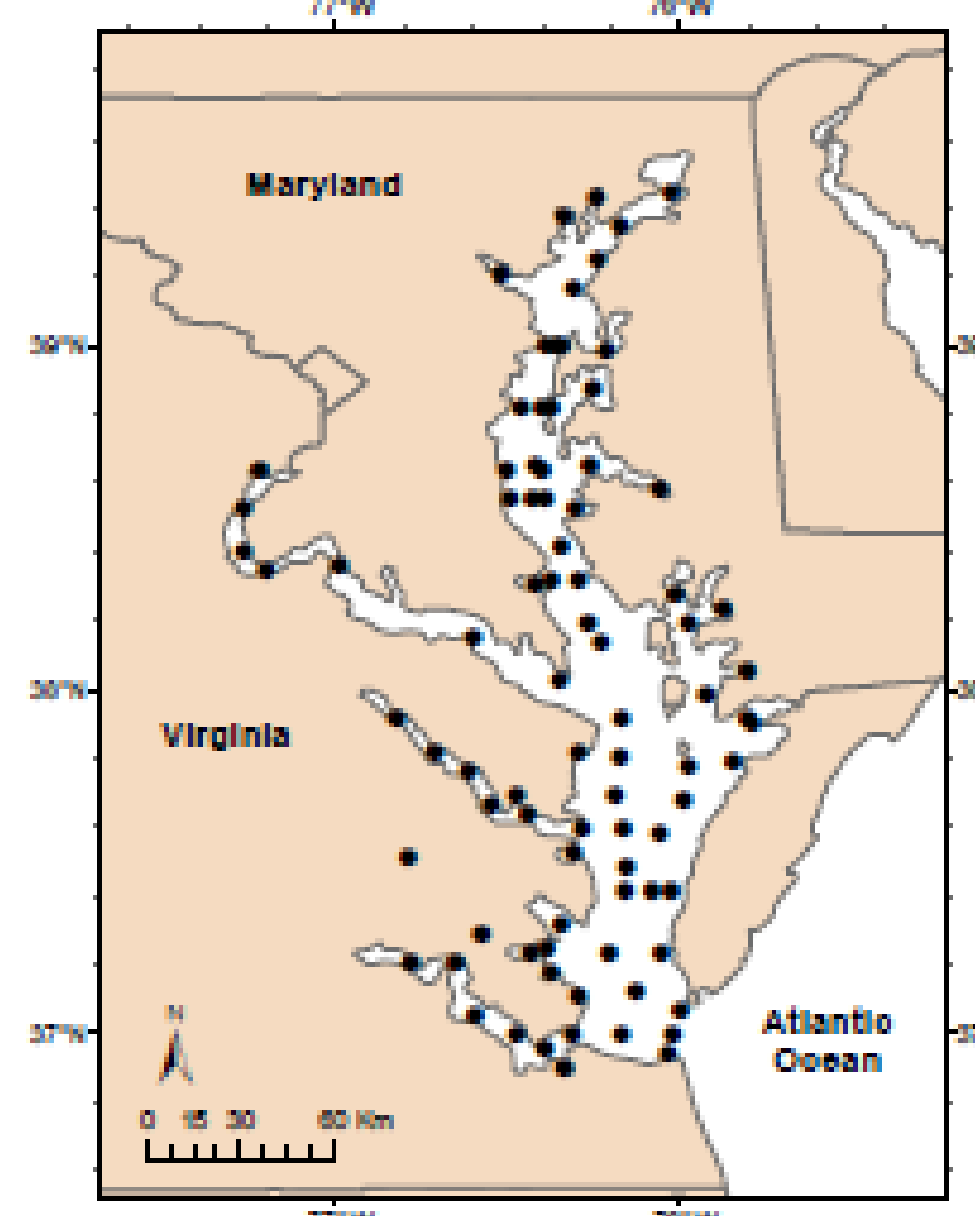


Figure 2. Matchup locations in Chesapeake Bay

Statistical and machine learning models in this study

- Classification and Regression Tree (CART)
- Bayesian Additive Regression Tree (BART)
- Random Forest (RF)
- Generalized Additive Model (GAM)
- Generalized Linear Model (GLM)
- Multivariate Additive Regression Spline (MARS)
- Neural Network (NN)
- Support Vector Machine (SVM)

Optically based, single-band algorithm for Chesapeake Bay (O-2012)³

$$TSS = 3.8813(nLw(645))^3 - 13.822(nLw(645))^2 + 19.61(nLw(645))$$

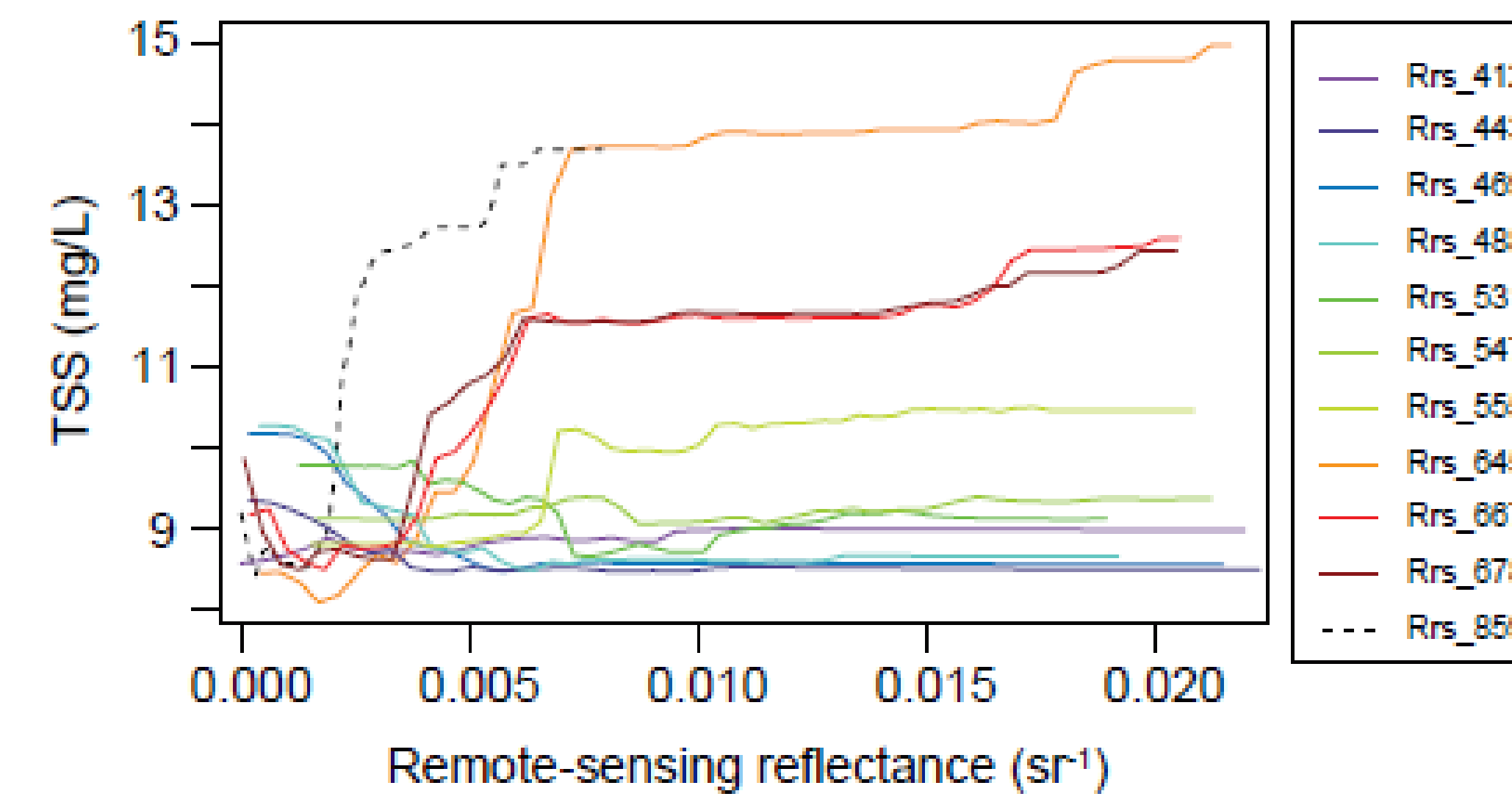


Figure 3. Partial dependence plots for the 11 MODIS bands used as predictor variables for TSS concentration in the Random Forest model before pruning.

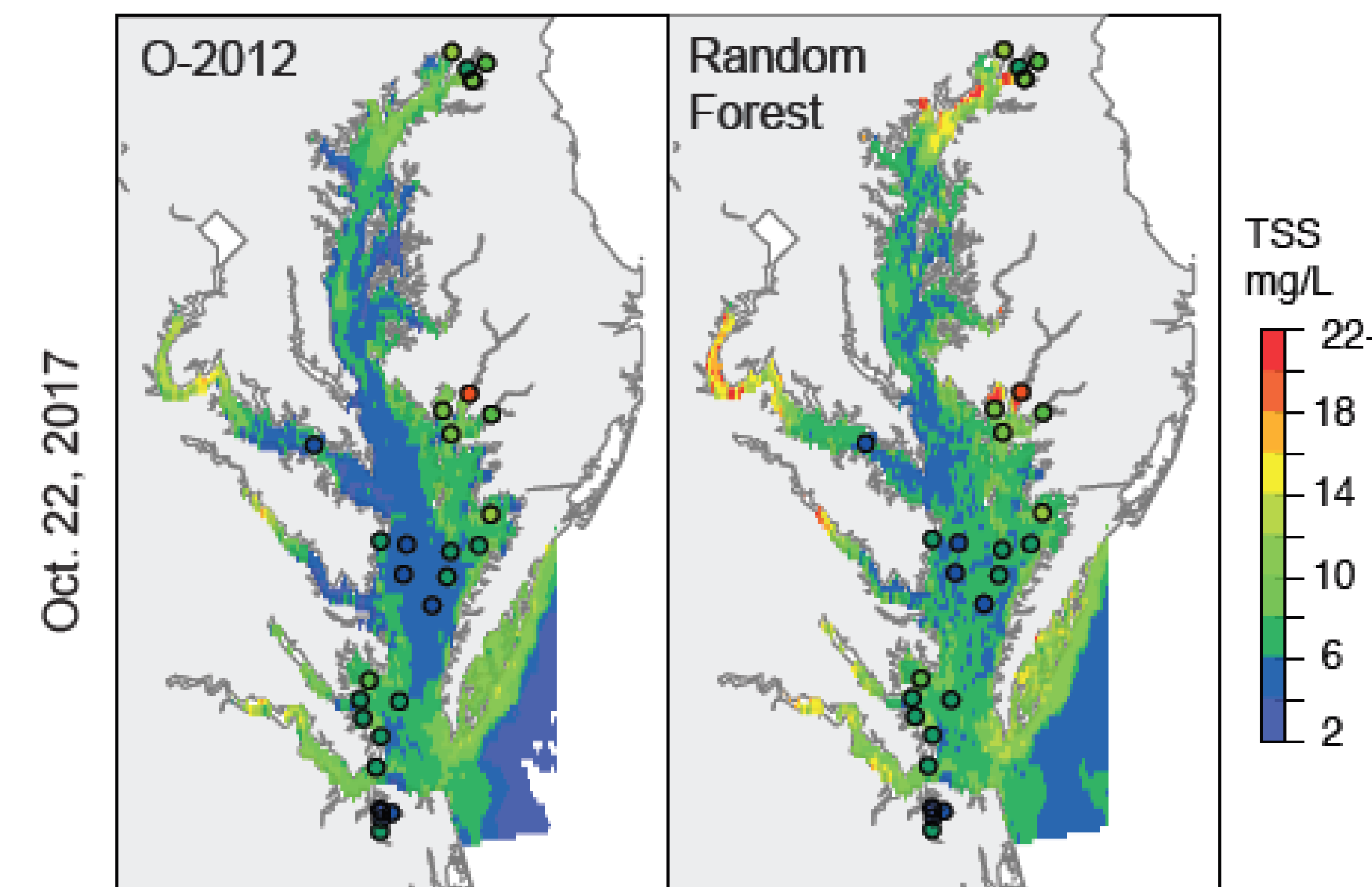


Figure 4. Comparison maps of MODIS TSS retrieval in Chesapeake Bay using the pruned Random Forest model and O-2012 algorithm for October 20, 2017.

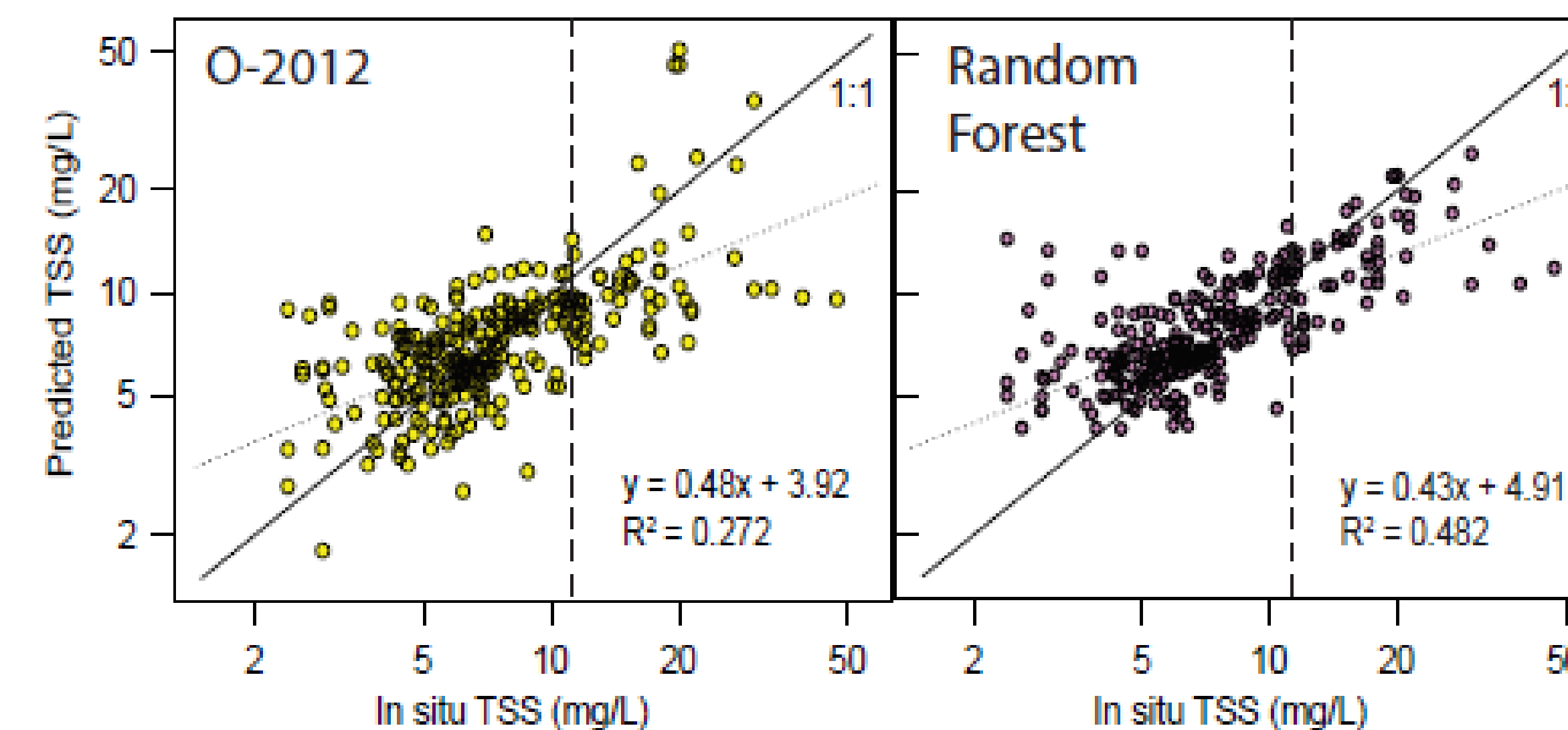


Figure 5. One-to-one regressions of CBP *in situ* measured versus MODIS satellite-derived TSS from single-band O-2012 algorithm and pruned multi-band Random Forest model.

Table 2. MAE, MSE, and RMSE for multi-band statistical and machine learning models (black headers) and single-band optical algorithm (blue header) on the 20% holdout validation dataset.

| | RF | GAM | GLM | NN | MARS | CART | BART | SVM | O-2012 |
|------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| MAE | 2.42 | 2.64 | 2.69 | 2.76 | 2.74 | 2.87 | 2.73 | 2.57 | 2.97 |
| MSE | 19.04 | 20.49 | 22.43 | 24.32 | 21.98 | 23.22 | 21.73 | 22.19 | 31.44 |
| RMSE | 4.36 | 4.53 | 4.74 | 4.93 | 4.69 | 4.82 | 4.66 | 4.71 | 5.61 |

O-2012(fit): Single-band polynomial with updated coefficients using our training data

$$TSS = 19.222(nLw(645))^3 - 10.293(nLw(645))^2 + 115.539(nLw(645)) + 8.467$$

Table 3. MAE, MSE, and RMSE for pruned RF model, (RF), Random Forest model using only 645-nm (RF(645)), O-2012 algorithm as published, and (O-2012(fit)) for the holdout validation dataset.

| Model | MAE | MSE | RMSE |
|-------------|------|-------|------|
| RF | 2.38 | 18.46 | 4.30 |
| RF(645) | 2.76 | 20.81 | 4.56 |
| O-2012 | 2.97 | 31.44 | 5.61 |
| O-2012(fit) | 2.71 | 21.77 | 4.67 |

Table 4. MAE, MSE, and RMSE for the RF model and O-2012 algorithm for holdout validation dataset above and below the 80th percentile (where TSS = 11.3 mg/L).

| | Above 80 th percentile | | | Below 80 th percentile | | |
|------|-----------------------------------|--------|---------|-----------------------------------|--------|---------|
| | RF | O-2012 | p-value | RF | O-2012 | p-value |
| MAE | 5.24 | 8.33 | 0.0006 | 1.7 | 1.62 | 0.3858 |
| MSE | 70.24 | 137.07 | 0.0095 | 6.06 | 4.67 | 0.0972 |
| RMSE | 8.38 | 11.71 | - | 2.46 | 2.16 | - |

Conclusions

- 1) Random Forest (RF) performs best of eight statistical models (Table 2). RF also performs better than optically based, single-band O-2012 algorithm on same holdout validation data (Table 2, Figure 5).
- 2) Machine learning approach in itself shows improvement over polynomial method using same MODIS band (Table 3).
- 3) RF outperforms the O-2012 on top 20th percentile of holdout data (TSS > 11.3 mg/L), particularly near-shore and in tributaries. However, the O-2012 performs better on the lower 80th percentile seen in the Mainstem (Table 4, Figure 4-5).
- 4) O-2012 is valuable tool for remotely retrieving TSS in the Chesapeake Bay for general environmental applications. Multi-band machine learning model could be useful where more precision in higher TSS waters needed.
- 5) Algorithms using red/NIR bands have high spatial resolution on MODIS (250 m). Multi-band MODIS models limited to 1 km resolution, but newer ocean color sensors provide higher resolution in all bands.

Future Work

Use RF model for TSS estimation in application where precision in higher TSS waters may be crucial. *Vibrio parahaemolyticus* is a naturally occurring bacteria in Chesapeake Bay waters that causes gastroenteritis when ingested in high concentration during raw oyster consumption. Studies have shown that suspended particles may be a useful environmental indicator for *Vibrio parahaemolyticus* to include in models that typically only use sea surface temperature and salinity.

References

1. DeLuca, N.M., Zaitchik, B.F., Curriero, F.C., 2018. Can multispectral information improve remotely sensed estimates of total suspended solids? A statistical study in Chesapeake Bay. *Remote Sens.* 10, 1393.
2. Dorji, P., Fearn, P., Broomhall, M., 2016. A semi-analytical model for estimating total suspended sediment concentration in turbid coastal waters of Northern Western Australia using MODIS-Aqua 250 m data. *Remote Sens.* 8, 556.
3. Ondrusek, M., Stengel, E., Kinkade, C.S., Vogel, R.L., Keegstra, P., Hunter, C., Kim, C., 2012. The development of a new optical total suspended matter algorithm for the Chesapeake Bay. *Remote Sens. Environ.* 119, 243-254.

Geospatial surveillance of opioid overdose mortality in 11 US states using geographic information systems (GIS) and Twitter API platform

Author: Randall Fowler (MS GIS, MS Ecology / Evolutionary Biology), PI: Dr. Carlos Castillo-Salgado, Association: JHSPH Global Public Health Observatory

Introduction:

The opioid epidemic in North America (USA and Canada) has reached dramatic levels. Opioid addiction has increased primarily due to the overprescribing of opioid medications for pain (CDC, Gomes et al., 1). Finding new ways to analyze how this epidemic is unfolding will provide public health personnel with a clearer understanding for what measures to implement to address this insidious disease. Here, I determine whether the news coverage of the opioid epidemic is in itself contributing to the pervasiveness of the epidemic by creating an opioid epidemic “meme” that adolescents respond to. The opioid epidemic “meme’s” effectiveness at promoting adolescent use of opioids is first measured by determining the meme’s prevalence in social media (Twitter) and then comparing that prevalence to opioid overdose mortality age-adjusted rates using standard regression analysis.

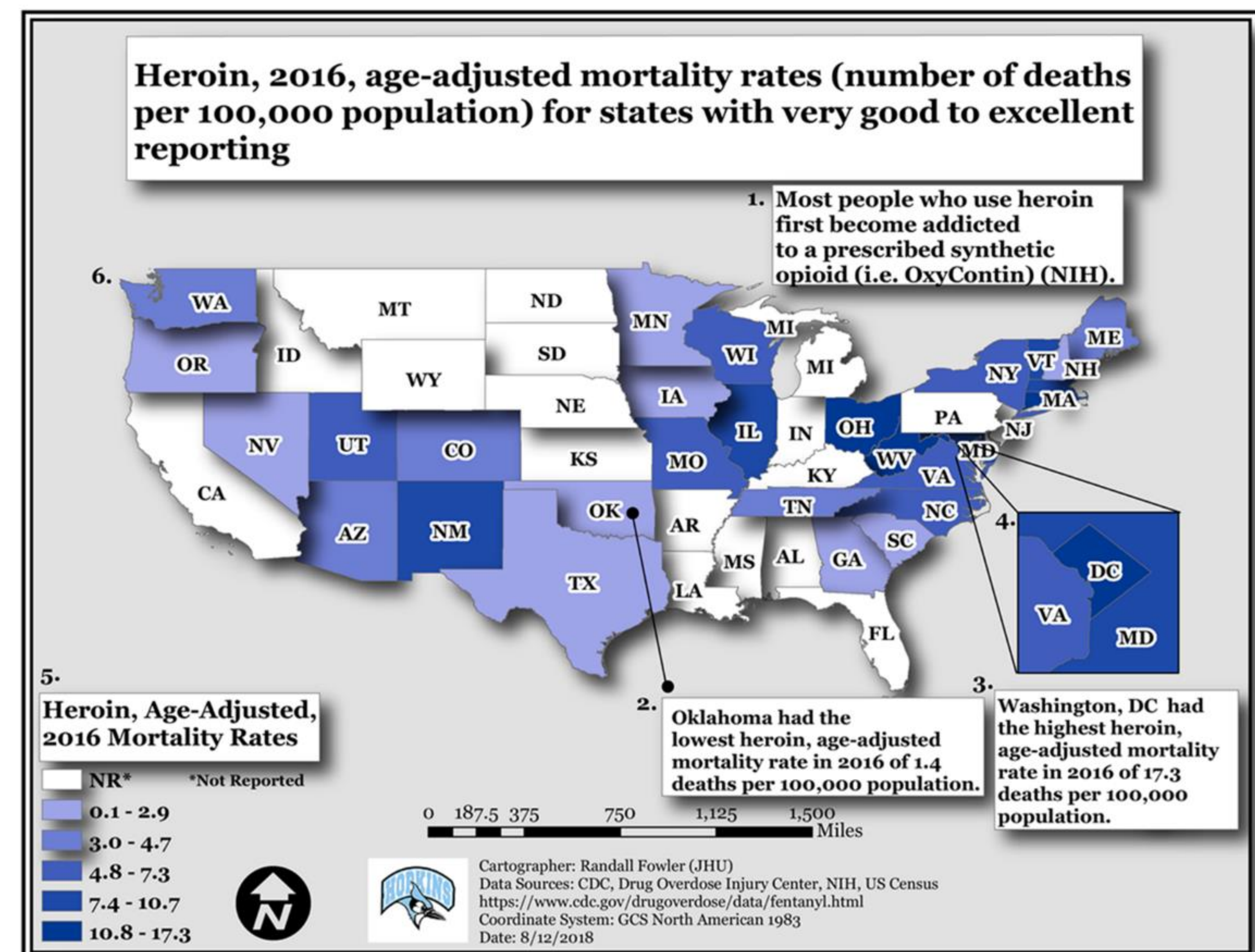
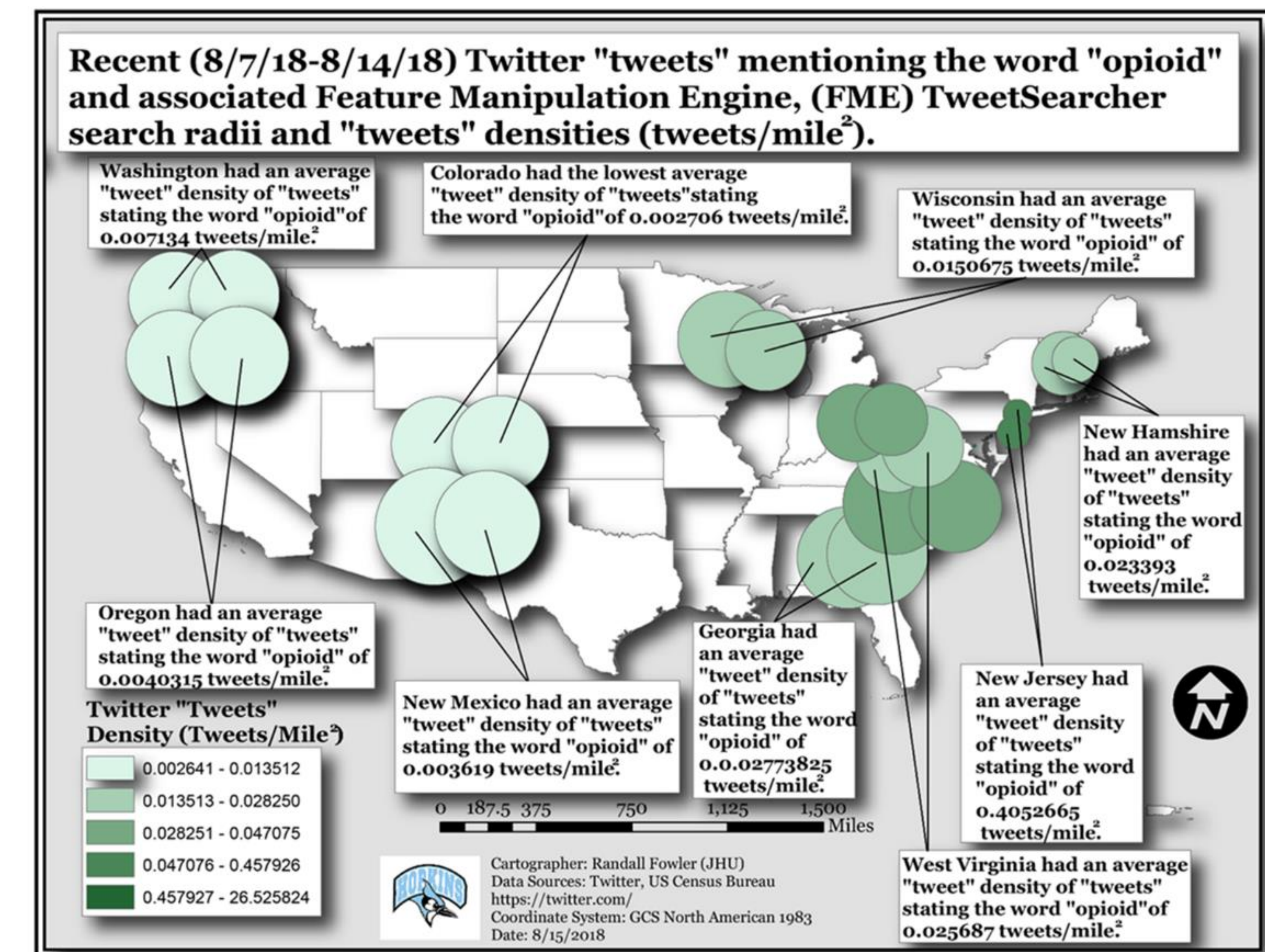
H1: There is a statistically significant, positive relationship between “tweet” densities mentioning the word “opioid” and opioid overdose mortality age-adjusted rates in 11 states.

Methods:

Overdose vital statistics data for 11 states / districts (CO, NM, OR, WA, WI, NH, WV, GA, NC, OH, NJ and DC) for 2016 was downloaded from Centers for Disease Control (CDC) Drug Overdose Injury Center (<https://www.cdc.gov/drugoverdose/>). The shapefile containing state boundary geometries was downloaded from US Census Bureau (<https://www.census.gov/geo/maps-data/data/tiger-cart-boundary.html>) at 500 km resolution. Centroids were established for the states in question in order to approximate the appropriate search radii for “tweets”. Additional features extrapolated from these centroids representing search radii (circles) were also digitized in an edit session. These radii were recorded in a configuration file that FME uses to extract the “tweets”. The search radii were digitized so the entirety of states was encapsulated by the new features. FME (version: 2017.1) was then used to extract “tweets” mentioning the word “opioid” from Twitter API for the most recent week of “tweets” (8/7/18-8/14/18). “Tweet” counts mentioning the word “opioid” were then determined for the states in the aforementioned search period. “Tweet” densities were calculated. A regression analysis was then used to determine if there was a statistically significant relationship between “tweet” densities (“tweets” per mile²) mentioning the word opioid by state, and opioid overdose mortality age-adjusted rates provided for in the CDC vital statistics.

Results:

There was found to be a non-significant (though almost significant) association between “tweet” densities mentioning the word “opioid” and heroin overdose mortality age-adjusted rates in the 11 states / districts of CO, NM, OR, WA, WI, NH, WV, GA, NC, OH, and DC ($r^2 = 0.36$, $F = 4.54$, $p < 0.07$, $SE = 4.67$, $df = 9$).



Discussion:

The results show that conversations in social media may permeate the ambient environment from cyberspace, influencing people’s lives and in this study, their deaths. Addictive behaviors like excessive use of social media and opioids revolve around a hijacking of the reward system inside the brain; where excessive substance abuse yields ever diminishing, pleasurable returns (Harvard Medical School). The social environment people find themselves in heavily influences their decisions, leading to many types of behaviors including substance abuse / addiction.

Image Mining of Aquatic and Household Coordinates in Choma District, Zambia



Brendan F. Fries¹, Kelly M. Searle², Timothy M. Shields¹, William J. Moss^{1,3}, Douglas E. Norris³

¹Department of Epidemiology, Johns Hopkins Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, ²Division of Epidemiology and Community Health, University of Minnesota School of Public Health, Minneapolis, MN, USA, ³W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA



Data Harvesting of Sites for Malaria Transmission Studies and Modeling, Pansharpening, Segmentation, and Unsupervised ISO for Household Classification

The Johns Hopkins Malaria Research Institute (JHMRI) is working with local partners to develop strategies to eliminate malaria transmission in Choma District of Southern Province, Zambia. Around the town of Macha we are conducting multiple studies to model and evaluate effective interventions combat transmission and infection.

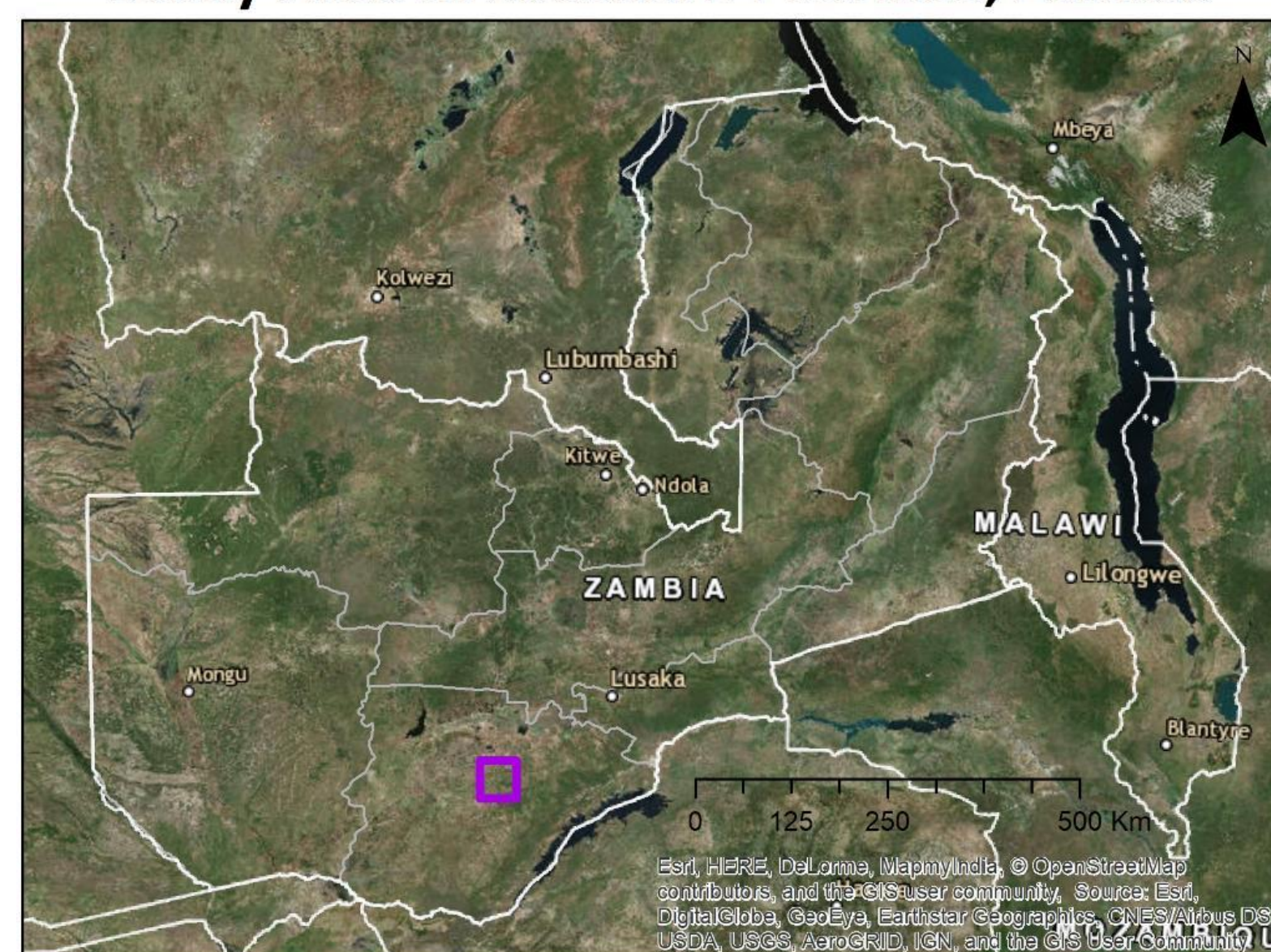


The Butterfly Tree - Malaria Prevention
<http://www.thebutterflytree.org.uk/pages/2014/malaria-prevention-2/>

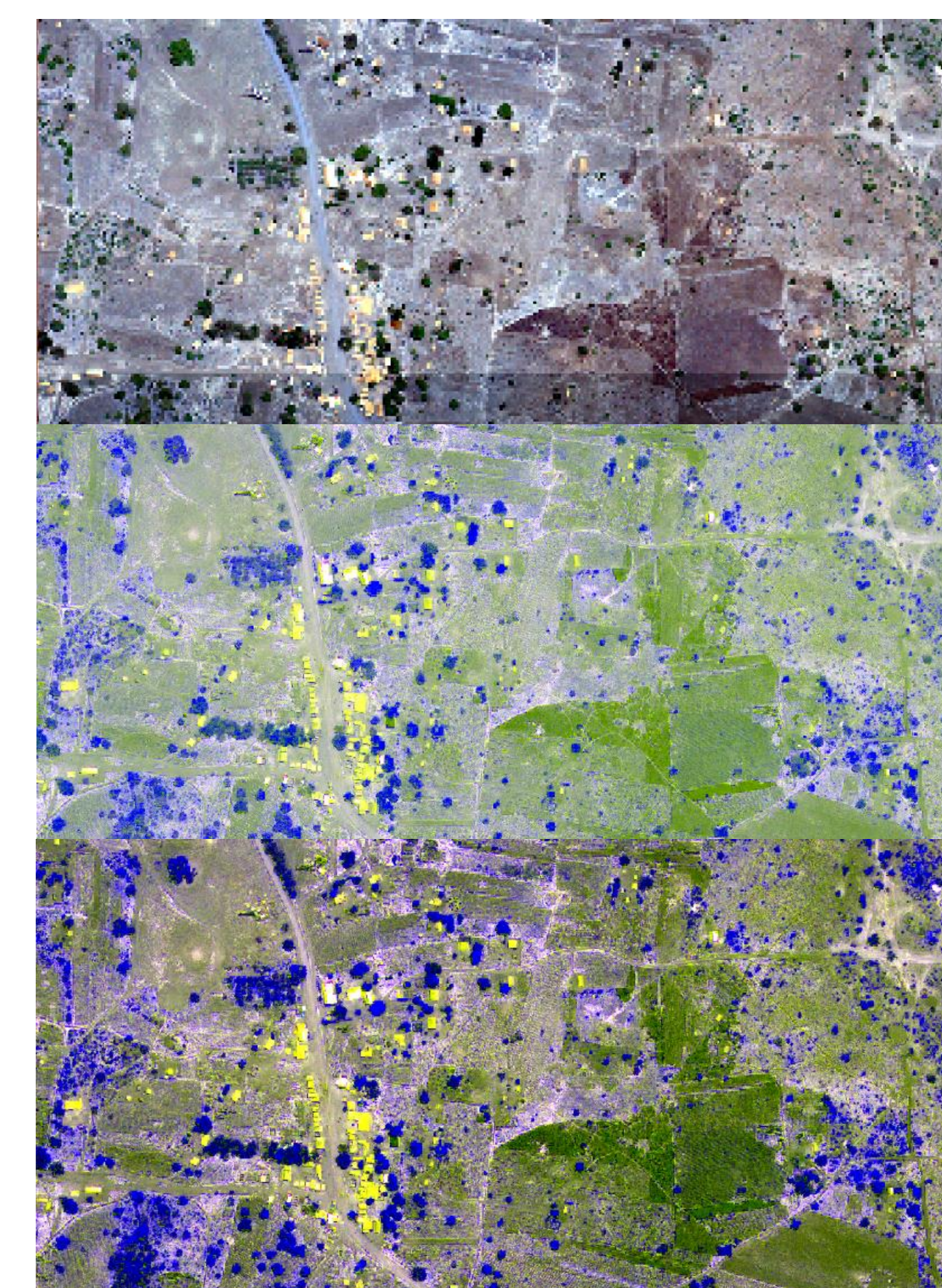


The Current - UCSB Computer Science and Film and Media Studies
<http://www.news.ucsb.edu/2011/01/13/20/ucsb-computer-science->

Study Area In Southern Province, Zambia



Modelling local transmission and conducting geostatistical analysis of malarial transmission dynamics require location data for several types of sites. Aquatic habitats hosting *Anopheles* mosquito breeding and households where human blood meals reside are both too numerous and expensive to manually geolocate. Thus we sought to mine satellite imagery of Macha and the surrounding

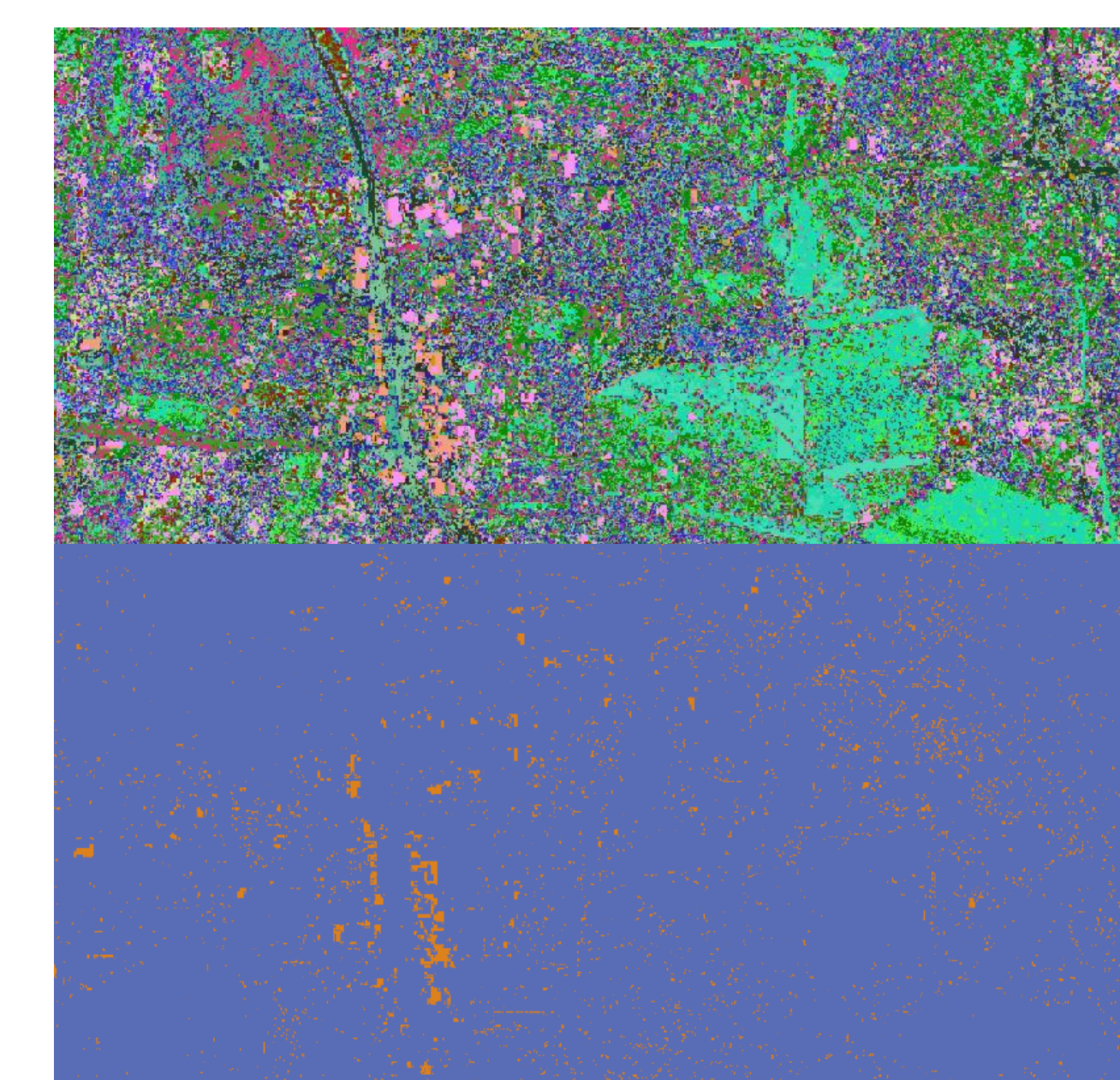


Pansharpening was used to fuse the higher-resolution panchromatic image to the multiband raster dataset in order to increase the resolution of the multiband raster dataset for the area covered to resolve the smallest household structures (3m).

We employed segmentation to group adjacent pixels that displayed similar color characteristics in the raster images as this to create a larger spectral distance between natural roofed structures and ground cover.

Next, we performed an unsupervised Iterative Self-Organizing Data Analysis Technique (ISODATA or ISO) classification on these prepared rasters. We selected several different parameters for the ISO Cluster analysis and ran a 25 class classification method.

The ISODATA algorithm used determines the characteristics of the natural groupings of cells in multidimensional attribute space. The clustering algorithm is a modified k-means clustering that features the merging of clusters if their separation distance in multispectral feature space is less than a user-specified value and pre-set cluster splitting rules. This method makes a large number of classes were found to represent the ideal number for our raster datasets and clusters of households. As the last step, the analyst visually chose the single ISODATA class that had the best spectral fit for households and the image was reclassified on a binary selection of 1/24 classes.



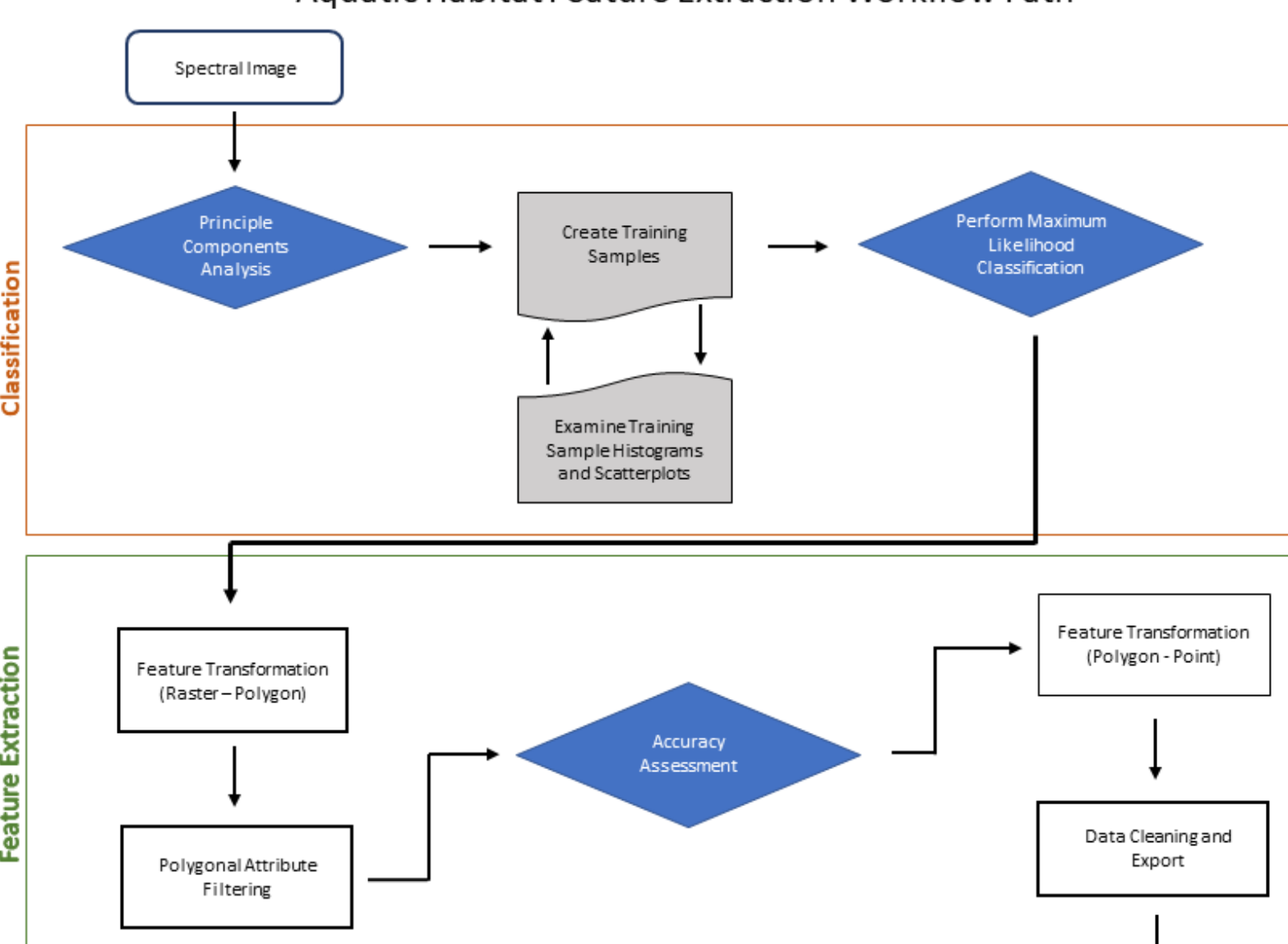
the dataset until specified results are obtained. The number of classes were found to represent the ideal number for our raster datasets and clusters of households. As the last step, the analyst visually chose the single ISODATA class that had the best spectral fit for households and the image was reclassified on a binary selection of 1/24 classes.

(Top L to Bottom R, Multispectral raster, pansharpened raster, segmented raster, ISO classified raster, binary raster)

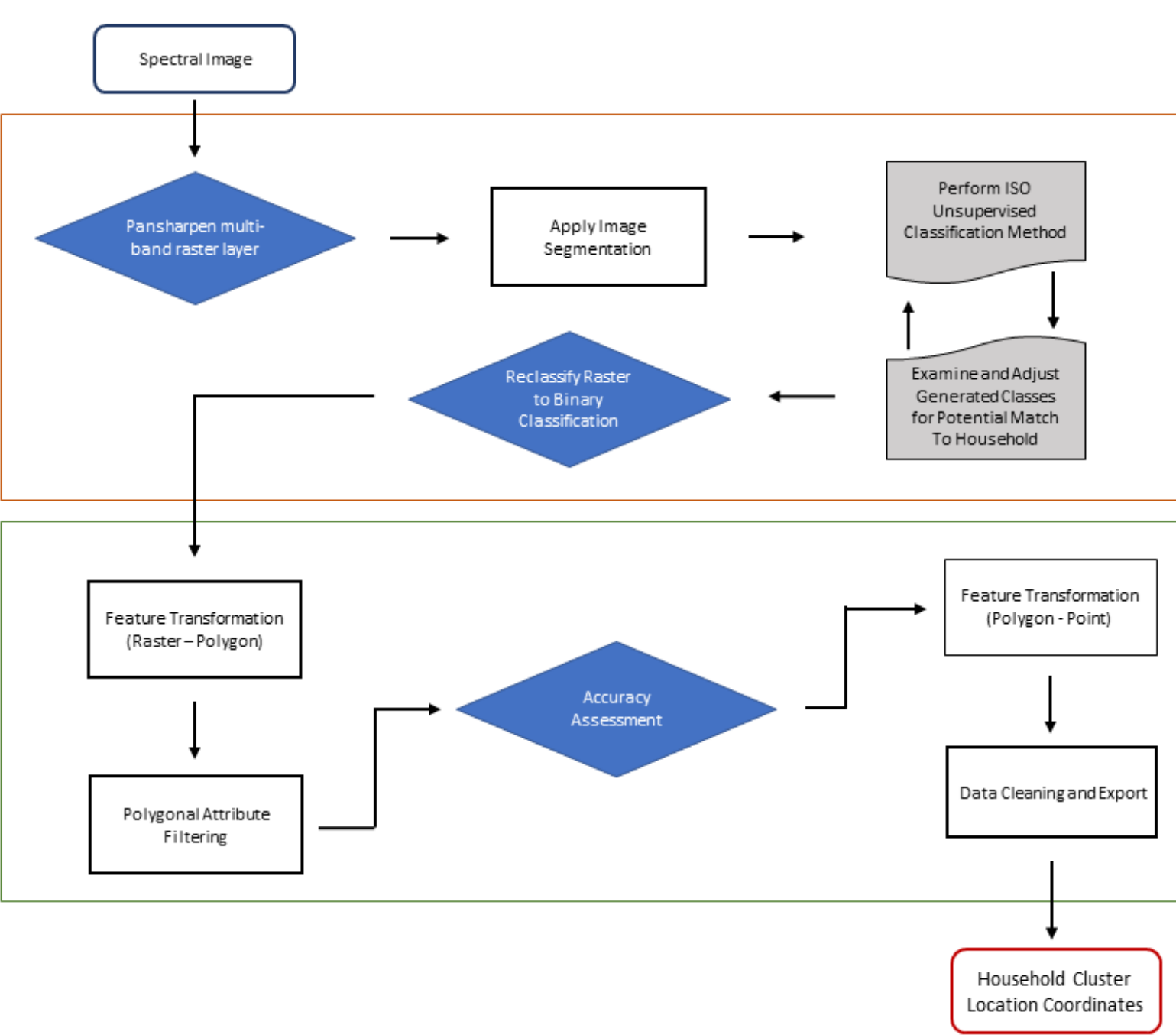
Extraction Model Process for Features of Interest

Starting with 2400 km² of 4-band 1.65m spatial resolution, 0.41m panchromatic resolution, imagery from the GeoEye-1 satellite, we utilized ArcMap 10.6 and ArcGIS Pro 2.3.3 to conduct our GIS analyses and develop our model. We required two different approaches to harvest spatial coordinates for aquatic habitats and potential households because of the high-resolution/low-bands necessary to resolve features as small as 3m across. For this reason we created two different feature coordinate extraction models.

Aquatic Habitat Feature Extraction Workflow Path



Potential Household Structure Extraction Workflow Path



Analysis of PCA of all bands by a binary supervised Maximum Likelihood Classification (MLC) using training samples created by the analyst. The classification stage of the Host Household model applied pansharpening of the multispectral with the panchromatic raster followed by a 25-class Unsupervised Classification which was reclassified on a binary raster based on analyst class selection.

Both models shared techniques for feature extraction once a binary raster of potential sites was developed. This consists of a transformation from raster to polygon, and then filtering of polygonal shape based on empirical data. Following an accuracy assessment, another transformation to point features was done before geolocation of X and Y coordinates and any excess or redundant fields

Accuracy Assessments of Model Output Features

| Classifier | Aquatic Habitat | Host Household |
|----------------------|-----------------|----------------|
| Sensitivity | 94.92% | 83.53% |
| Specificity | 91.36% | 94.46% |
| PPV | 88.89% | 87.51% |
| NPV | 96.10% | 97.62% |
| F ₁ score | 0.9181 | 0.8547 |

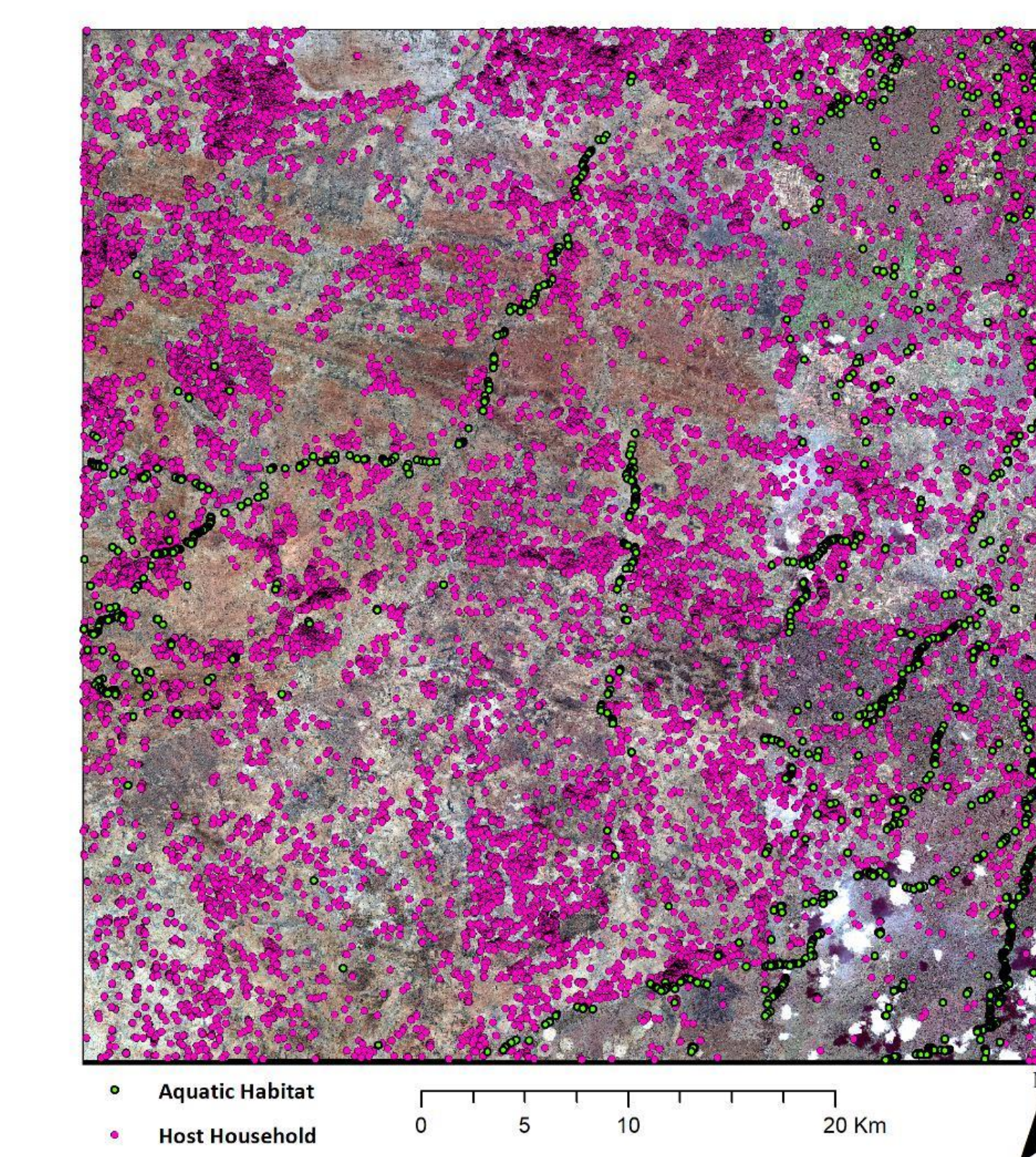
We utilized a confusion matrix of points across an area of 350 km² in order to assess the accuracy of our geolocation models. This was developed using manual coordinate gathering by two different human analysts.

The confusion matrix actual class for the aquatic surface features was composed of 130 pre-identified positive point features and 670 randomly generated negative points. Matching consisted of coordinates falling within 15m of each other. The confusion matrix actual class for household cluster features was composed of 5912 pre-identified positive point features and 2679 randomly generated negative points. Matching consisted of falling with same household cluster (100m SE +/- 15 m).

True specificity for the aquatic and household models was possibly higher due to limited randomly generated negative points and does not represent the ratio of total negative pixels to positive across the imagery (>0.01%)

Harvesting of remote sensing imagery data resulted in enumeration of:

- **1,702 aquatic habitat coordinate locations**
- **27,548 household coordinate locations**



Acknowledgements:

We would like to thank the study team at Macha Research Trust, Choma. We are grateful to the community of Choma District for their cooperation.

Funding:

This research is supported in part by the Bloomberg Philanthropies and the Johns Hopkins Malaria Research Institute, and the NIH-sponsored Southern and Central Africa ICEMR 2U19AI089680. ArcGIS software license was provided by the GIS & Data Services of the Sheridan Libraries at Johns Hopkins University.

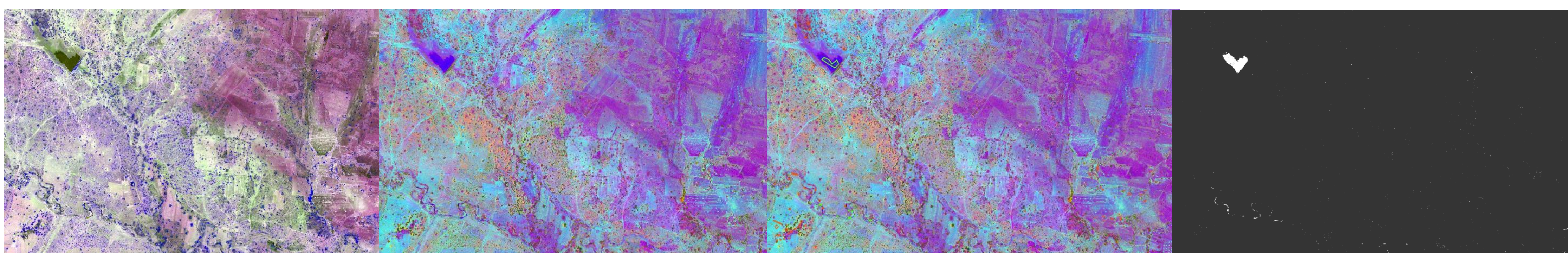
Conclusions

We found that the feature extraction models we developed were able to identify the features of interest for malaria transmission with a high degree of accuracy. The F-score for the Host Household model at 0.85 is comparable to the accuracy of contemporary building identification models that have been developed for natural material roofed structures.

We observed significant increase in the number of aquatic habitats present during the wet seasons which corresponds to the increase in seasonal transmission of malaria observed in the region. Across the study area household density ratios varied from <1 to 350 per km² with an average of 14 per km².

The resulting coordinate locations from our image mining of this spatial imagery will be used for simulation and evaluation of an agent-based transmission model of malaria in the Choma region.

PCA and Supervised MLC for Surface Aquatic Classifier



(L to R, multispectral raster, PCA raster, training samples, binary raster)
We ran a PCA on our 4-band imagery and chose a 3 band output which allowed us to produce a 3 band RGB image that retained ~95% of data and eliminating redundancy thus making our classification methods simpler. PCA assigns data to eigenvectors based on variability levels, ranks these on eigenvalue and in doing so reduces dimensionality which results in a simplified dataset. Our *principle components* output raster was then used to collect our two-class water/not water training samples by

These training samples were fed into an MLC which produced the binary classification in black and white. The MLC classified the input raster using a *signature* based on our provided training samples.

MLC functions on the assumption of cells in each class sample in the multidimensional space being normally distributed and Bayes' theorem. The tool considers both the variances and covariances of the class signatures when assigning each cell to one of the classes represented in the signature file.

Potential of the GIS enabled digital data collection tool for insecticide-treated nets mass campaigns in Nigeria: Lessons from a formative study

Youngji Jo¹, Nathan Barthe², Elizabeth Stierman², Kathryn Clifton², Sonachi Ezeiru³, Diwe Ekweremadu³, Nnaemeka Onugu³, Imran Chishtie³, Zainab Ali³, Elijah Egwu³, Ochayi Akoh³, Kingsley Godson³, Orkan Uzunyayla⁴, Suzanne Van Hulle²

1. Johns Hopkins Bloomberg School of Public Health | 2. Catholic Relief Services (Headquarter) | 3. Catholic Relief Services (Nigeria office) | 4. Red Rose

BACKGROUND

- The combination of mobile, coupled with spatial, technology applications and high-resolution satellite maps has created new avenues to increase capacity of monitoring field-based activities in real-time.
- In Nigeria, an information communication and technology (ICT) platform was piloted to facilitate planning and monitoring during the mass distribution of long-lasting insecticide-treated nets (LLINs) in the **Oyun** local government area (LGA) in 2017.

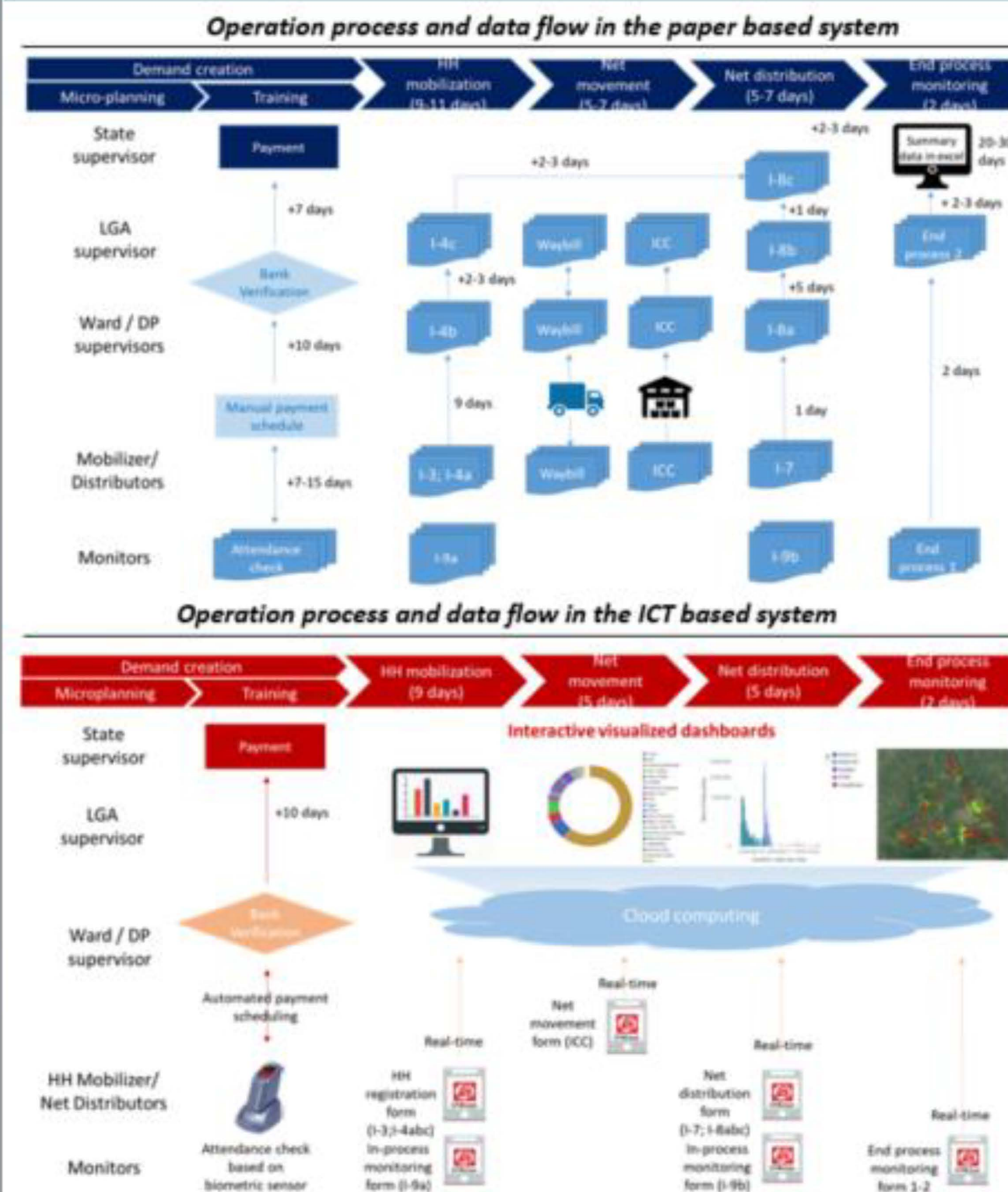
OBJECTIVES

- The formative study aims to identify major challenges and resolution approaches in the current paper-based system and to discuss how the ICT based systems can improve the management, monitoring and evaluation of LLIN mass campaigns in Nigeria.

METHODS

- Catholic Relief Services (CRS) implemented malaria LLIN replacement campaigns between April and November 2017 in the six states in Southeast Nigeria (**Kwara**, **Ondo**, **Imo**, **Edo**, **Osun**, and **Adamawa**).
- The ICT system was implemented in one local government area (LGA) as a pilot from April to May 2017 in **Kwara** State. The paper system was implemented, from July to August 2017 in the entire **Edo** State.
- We conducted qualitative and observational study to understand technical features and operational process of paper and ICT based systems based on the documented operation manuals, field observations, and informant interviews. In addition, we documented daily review meeting discussions during the campaign from six LGAs in **Edo** State.
- We categorized the major campaign implementation issues into logistic, technical, and demand creation issues.
- Overall campaign operation processes are similar but monitoring indicators, data collection and management systems differ between the paper and ICT groups.
- These differences lead to different data outputs and implications to efficiency, transparency, accountability of management decision making in LLIN mass campaign.

METHODS



RESULTS

Figure 1. Mapped household registrations by a worker



- The map shows the movement of a staff for the household registration (net cards issuance). The dots show the number of the net cards issued in the households.
- The data is linked with an individual staff ID and specific time/location/numbers of the net cards issued, and individual net card's QR code numbers.
- This allows to reduce any fraudulent net cards issuances, identify any missed households, and ensure net cards redemptions during the net distribution campaign.

Figure 2. Mobilization density on the map

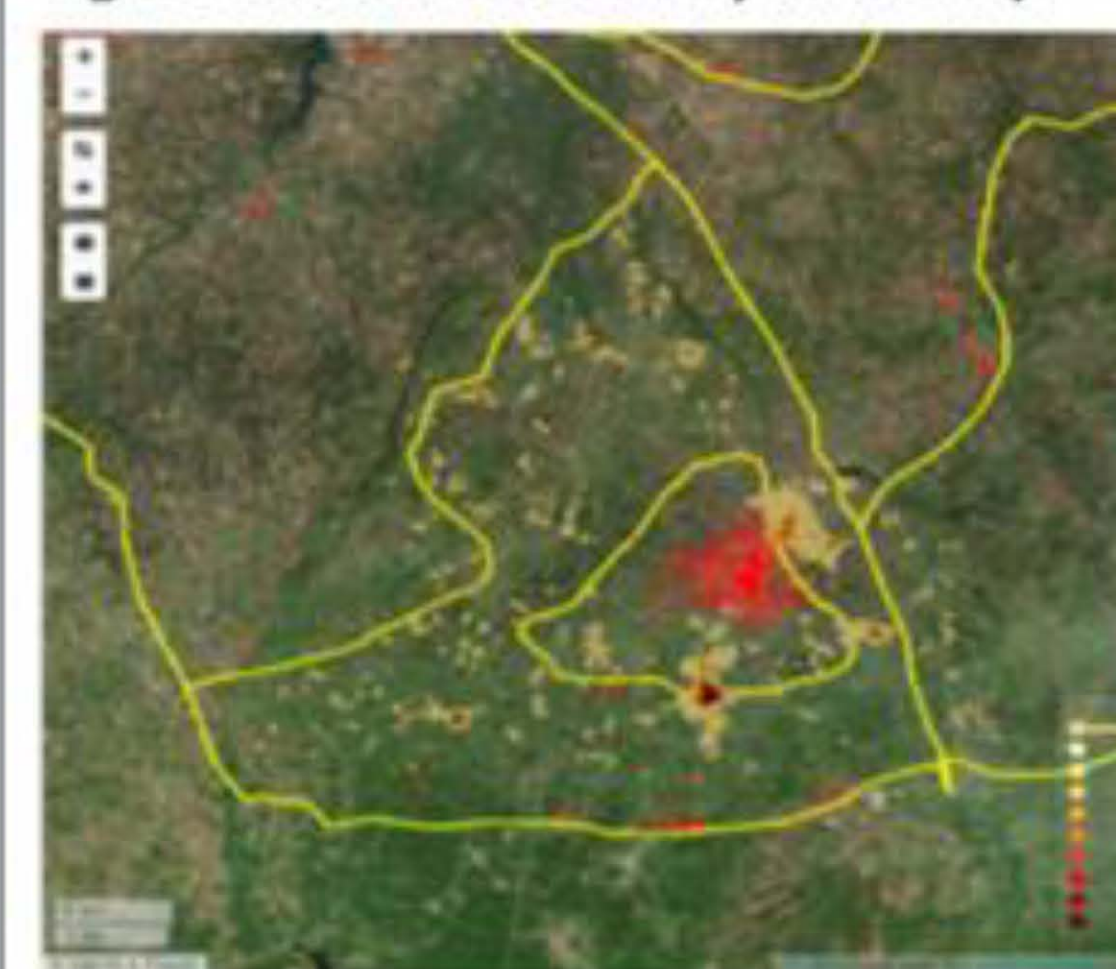


Figure 2: The bright red dots represent the locations of all households. The layer on top of it is made out of colored squares and that represents the coverage based on the amount of the net cards distributed.

Figure 3. Distribution Point distances

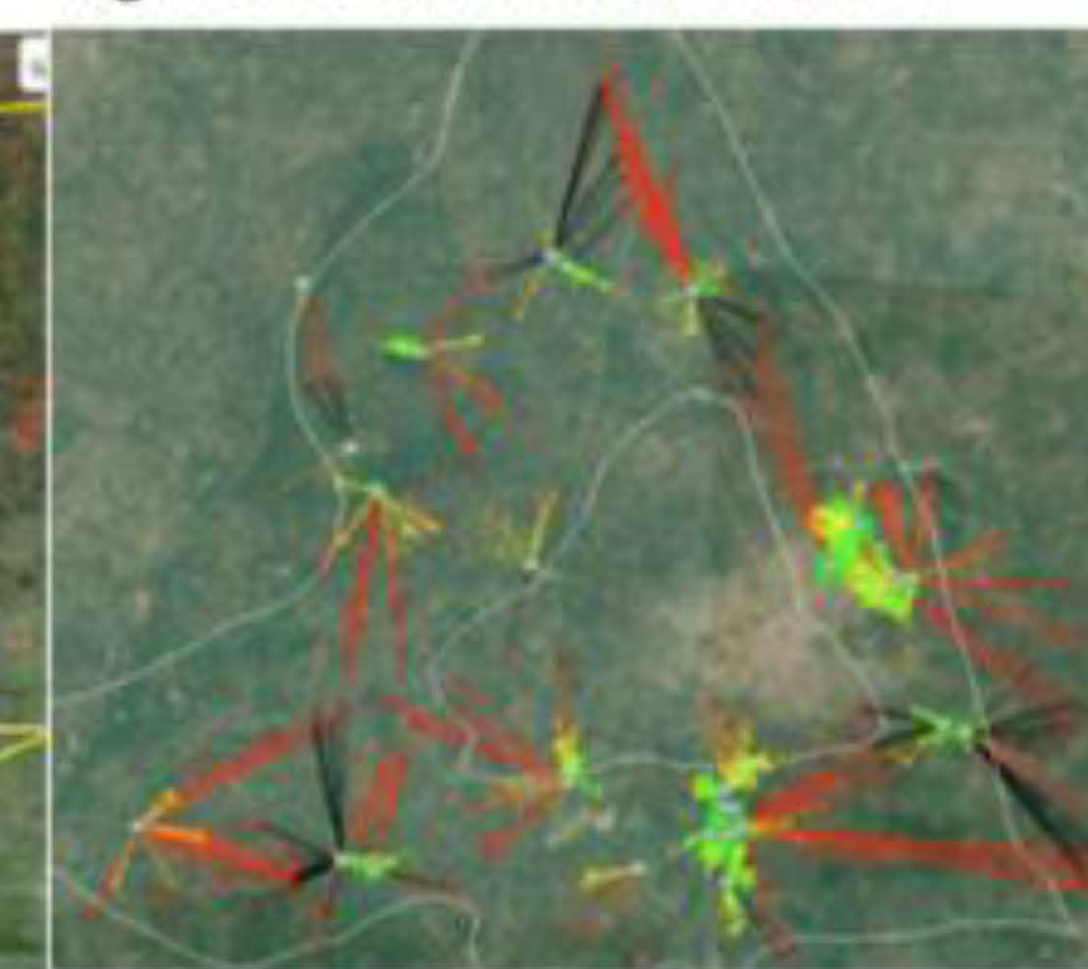


Figure 3: This map shows the location and distance of distribution points to households/towns. The allows to find the optimal (closest) distribution points based on the distance to the households and population density.

CONCLUSIONS

- We found that one of the biggest challenges in paper-based systems was the variability brought in by the human factor.
- Most issues are identified through staff reporting during review meetings or through oversight visits, and assurance is provided by occasional checks by supervisors to monitor attendance and staff performance.
- While the system requires a large volume of hard copies of reports or forms, information flow is one-way and not actionable.
- The ICT system can improve management and oversight through user-centered interface design, built-in data quality control logic flow/algorithms, workflow automation, and visualized interactive dashboards.
- While the ICT based system would have similar implementation challenges as those in the current paper-based system, it can better manage issues related to the data reporting system and processes, which comprises significant staff time and effort in the current paper-based practice.

Main Lessons

- ICT systems are likely to improve management and oversight for LLIN mass campaigns, through greater transparency and efficiency in the data collection and reporting process.
- ICT system holistically changes overall program process and dynamics—not simply expediting the current process—it requires developing new operation and training manuals and policies.
- By closing the feedback loop with actionable data, such a tool can facilitate planning, promote workforce performance, and enhance accountability.
- Our findings may provide useful insights for the development and use of technology in public health campaigns.

REFERENCES

- The Alliance for Malaria Prevention, *A toolkit for mass distribution campaigns to increase coverage and use of long-lasting insecticide-treated nets*, 2012.
- Teklehaimanot, A., J.D. Sachs, and C. Curtis, *Malaria control needs mass distribution of insecticidal bednets*. *Lancet*, 2007. 369(9580): p. 2143-2146.
- Paintain, L.S., et al., *Sustaining fragile gains: the need to maintain coverage with long-lasting insecticidal nets for malaria control and likely implications of not doing so*. *PLoS One*, 2013. 8(12): p. e83816.

Seasonal bionomics of malaria vectors in Kilwa and Kashobwe, Haut-Katanga Province, Democratic Republic of Congo



Tamaki Kobayashi¹, Thierry Bobanga², Solange Umesumbu³, Serena Schatz¹, Christine M. Jones¹, Jennifer C. Stevenson¹, William J. Moss¹, Douglas E. Norris¹ for the Southern and Central Africa International Centers of Excellence for Malaria Research
¹Johns Hopkins Bloomberg School of Public Health, ²Université Protestante au Congo, ³National Malaria Control Program



Introduction

The goal of the Southern and Central Africa International Centers of Excellence for Malaria Research (ICEMR) is to address critical research questions on barriers to malaria control and elimination in Zambia, Zimbabwe and the Democratic Republic of Congo (DRC). Of importance for regional malaria control is movement of parasites and vectors across international borders. Nchelenge District, one of the study sites of Southern and Central Africa ICEMR in Zambia, is located near the border with DRC. In this study, baseline data are presented on malaria vectors in the southern Haut-Katanga Province, DRC, which shares a border with Luapula Province in Zambia, along Lake Mweru and the Luapula River.

Methodology

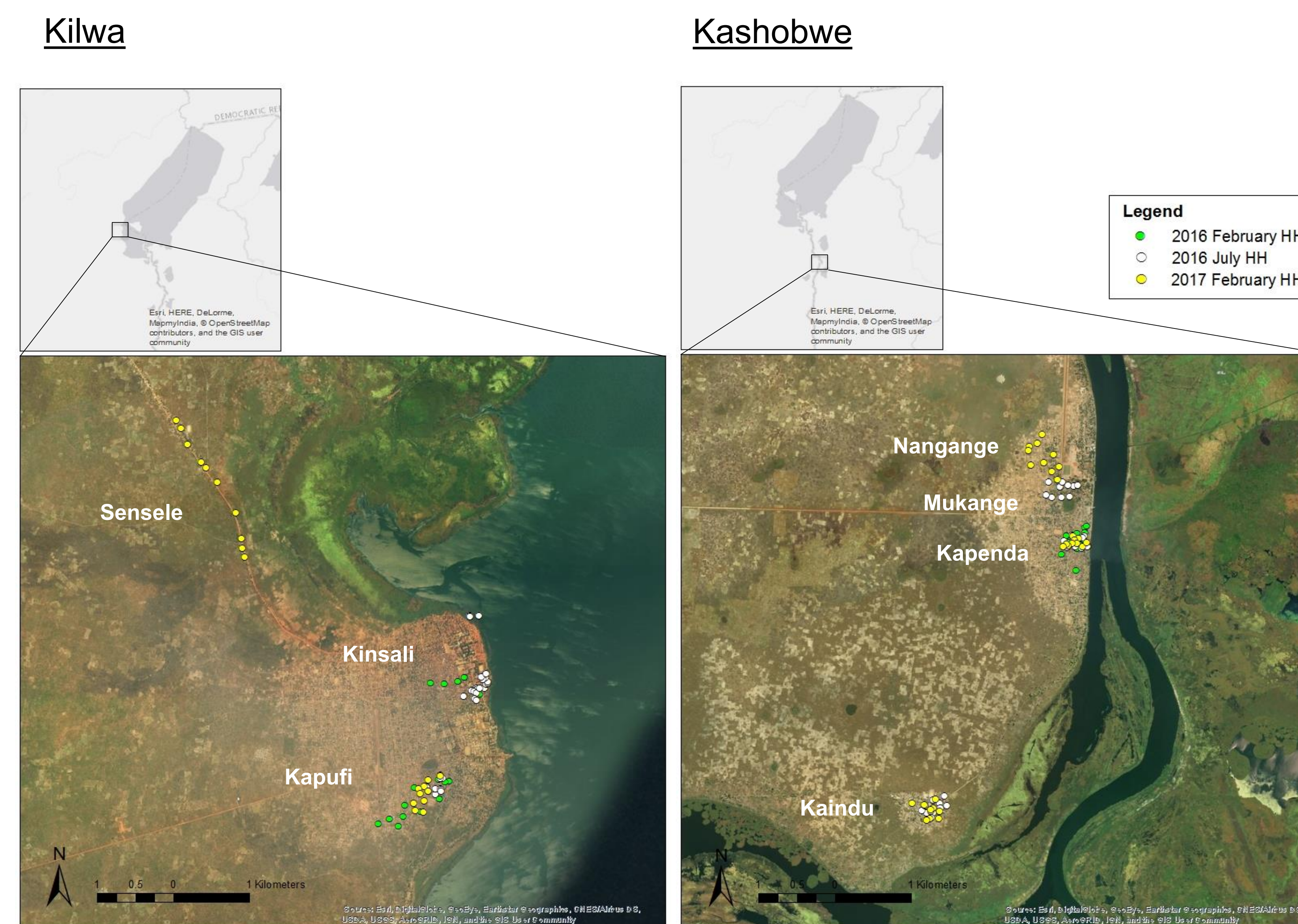
Three cross-sectional surveys were conducted in Kilwa and Kashobwe in Haut-Katanga Province in February (rainy season) 2016 and 2017 and July 2016 (dry season). Kilwa is located near Lake Mweru and Kashobwe is located near the Luapula River, providing abundant vector breeding sites throughout the year (Fig.1). Mosquitoes were collected using CDC light traps and pyrethroid spray catches (PSC). Species of female *Anopheles* mosquitoes were determined morphologically. For samples collected in 2016, species were further confirmed using PCR. Sources of blood meals were identified by PCR and sporozoite infection was identified by circumsporozoite protein (CSP) specific ELISA.

Figure 1 Location of Study Sites



From 7 villages, 38 households (HH), 60 HH and 50 HH were screened in February 2016, July 2016 and February 2017, respectively. In total, 822 female *Anopheles* mosquitoes were caught and analyzed: 189 mosquitoes in February 2016, 288 mosquitoes in July 2016, and 345 mosquitoes in February 2017. Blood meal PCR showed that the captured mosquitoes were highly anthropophilic.

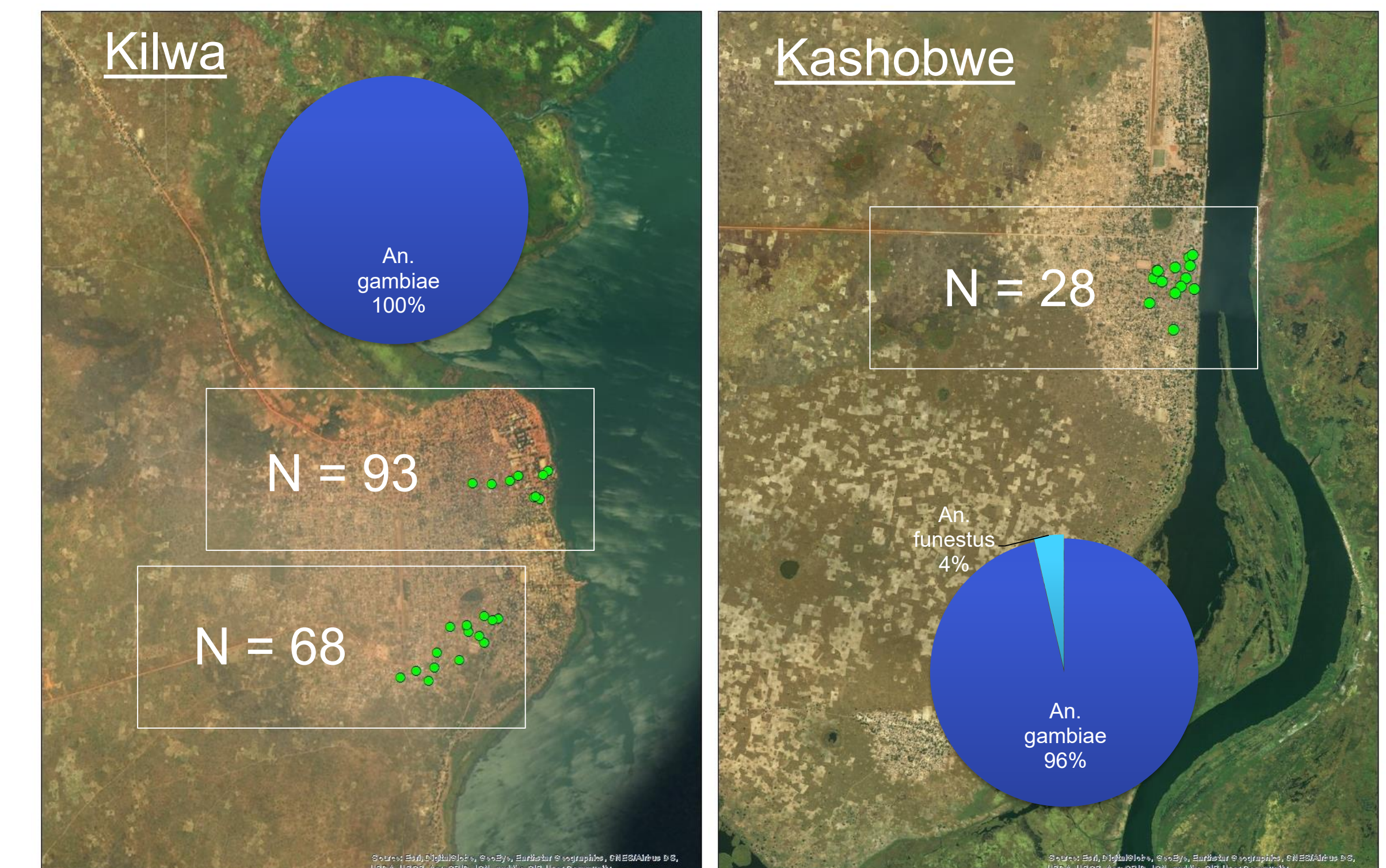
Figure 2 Location of households where mosquito collections were done



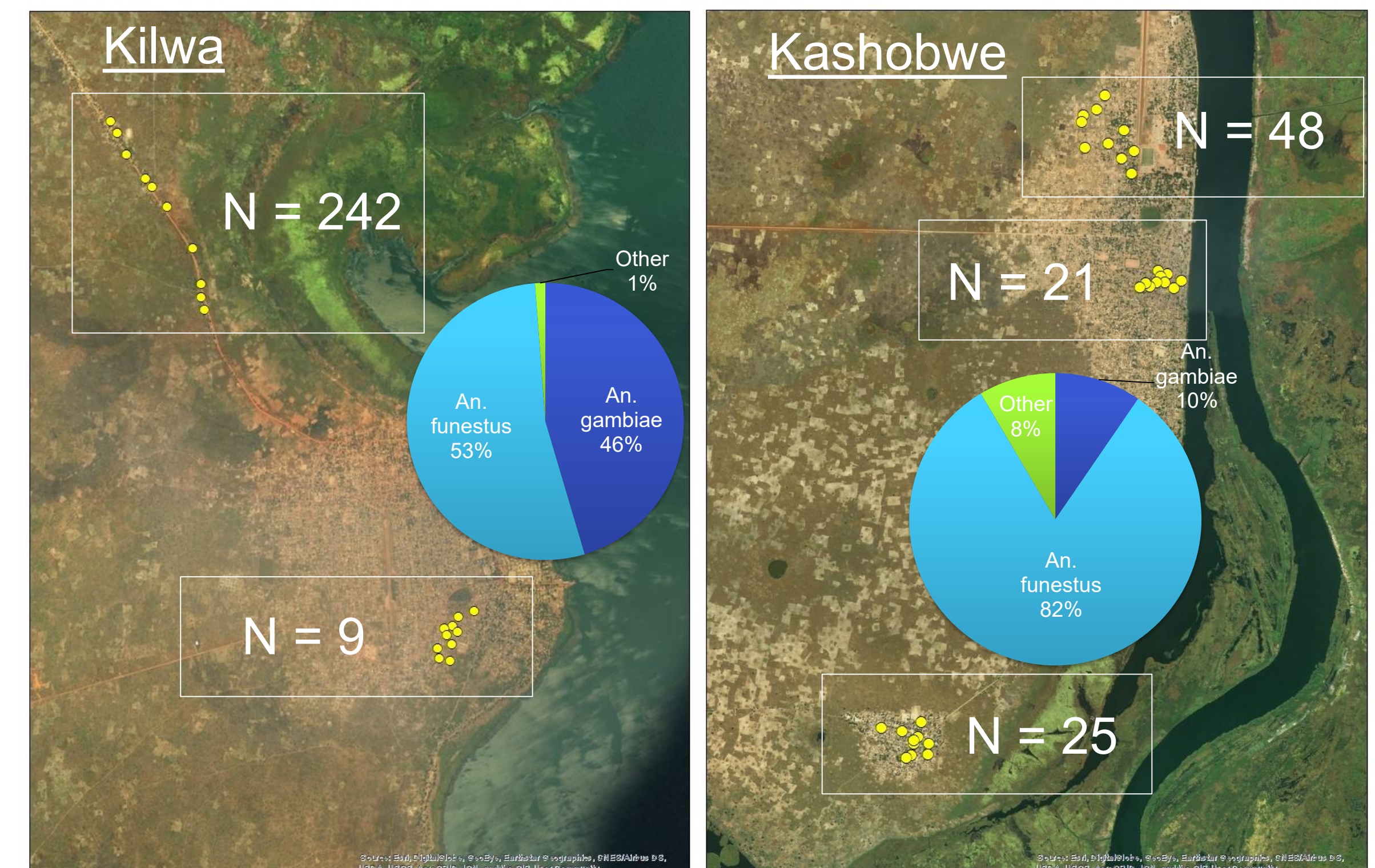
Results

Number of female *Anopheles* species caught in the rainy season

February 2016 (Confirmed by PCR)



February 2017 (Morphological identification only)



Other species include *An. gambiae*, *An. coustani*, *An. paludis*, *An. tenebrosus*.

Number of female *Anopheles* species caught in the dry season

July 2016 (Confirmed by PCR)

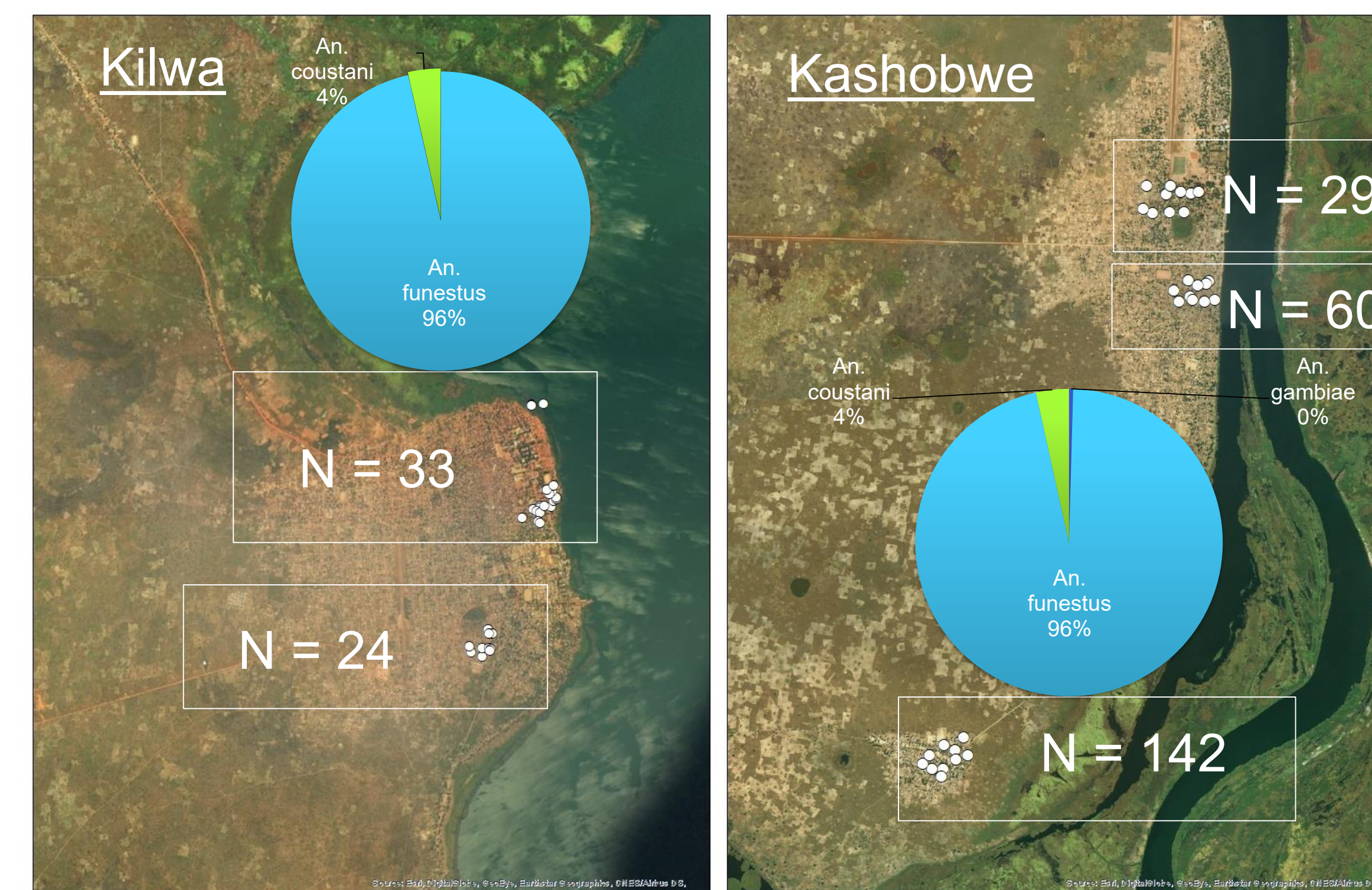


Table 1 Sporozoite infection rate by CSP ELISA

| | February, 2016 | July, 2016 | February, 2017 |
|--|----------------|----------------|----------------|
| Number of female <i>Anopheles</i> mosquitoes | 189 | 288 | 345 |
| Sporozoite infection rate: % (95% CI) | 0.5 (0.1, 2.9) | 2.4 (1.0, 4.9) | Not done |

Conclusions

This bionomic pattern of anopheline malaria vectors in these collections is similar to that observed in Nchelenge District, Zambia. Additional cross-sectional surveys are necessary to delineate the foraging and resting behaviors of mosquito vectors to guide malaria control strategies.

Background/Purpos

Understanding the role of environmental exposures in the development of SLE and their association with SLE activity may help identify modifiable risk factors and potential etiological mechanisms.

Cluster detection is an essential tool in public health which has the goal of detecting anomalous clusters of disease cases.

We performed a spatial-time cluster analysis of the Johns Hopkins Lupus cohort with the goal of identifying potential spatial-time clusters of SLE organ specific disease activity related to temperature and fine particulate matter pollution (PM2.5).

Materials and Methods

1261 patients who fulfill 4 of the 11 American College of Rheumatology classification criteria for SLE and who had recorded home addresses were included in the analysis.

Disease activity was expressed as Physician Global Estimate (PGA).

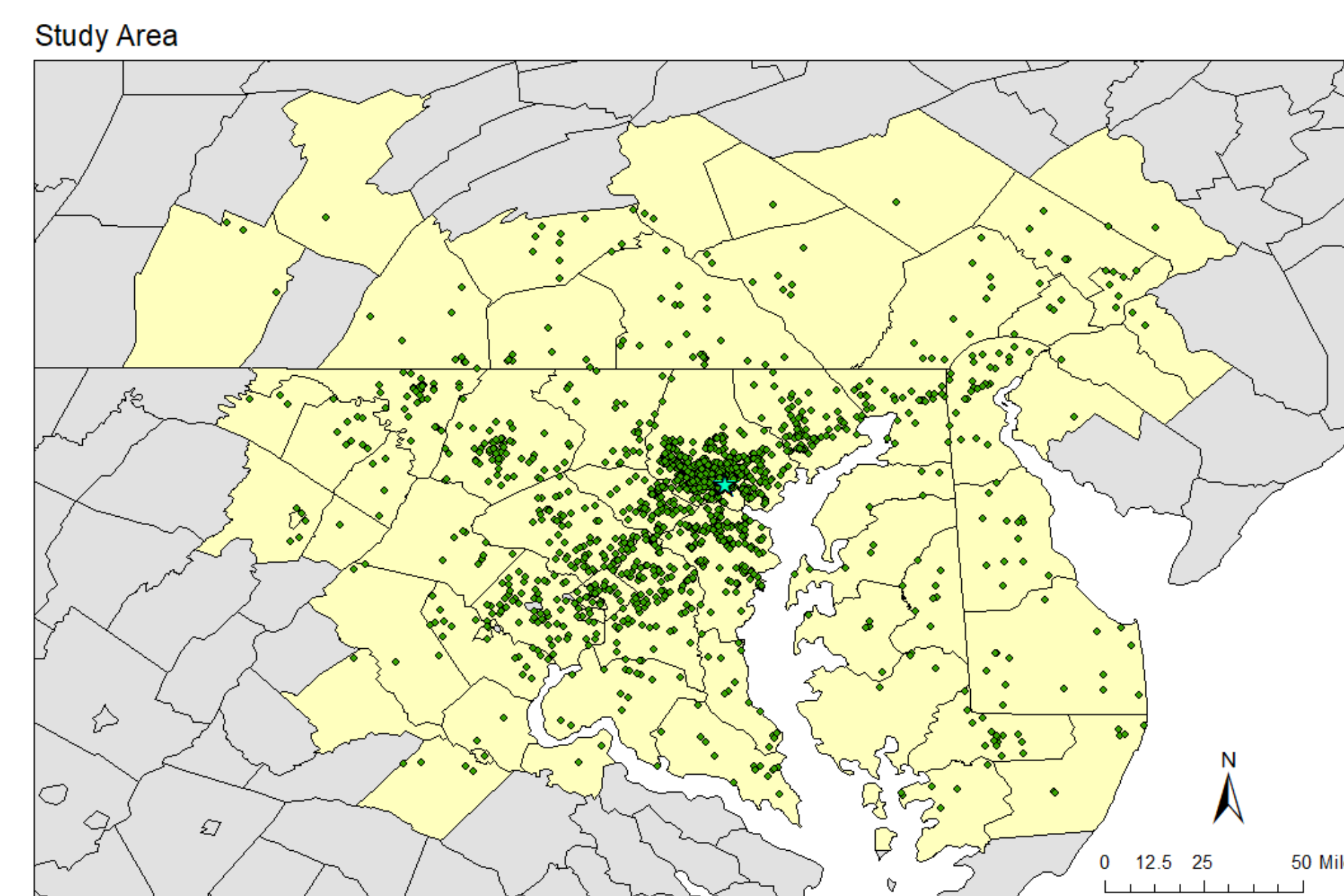
The area utilized in this analysis was a 350 kilometer radial buffer around the Johns Hopkins Lupus Center. This area was considered due to the high and consistent density of study participants. The data ranged from January 1999 to February 2009.

Average temperature and PM2.5 exposure over a period of 10 days prior to patient visit was obtained from the United States Environmental Protection Agency,

Univariate and multivariate models were built in order to study the association of these variables with lupus disease activity. The models were adjusted for age, sex, income, racial distribution, and rural vs. urban patient residence.

Spatiotemporal cluster detection was conducted using SaTScan v.9.4.4. software, considering monthly time intervals. Unadjusted cluster detection was first completed for all disease activity variables, and then repeated, adjusting for variables that were found to be significant in multivariate models.

Results



Legend
★ Johns Hopkins Lupus Center
● Lupus Cohort
■ Counties Included in the Analysis

Figure 1: Johns Hopkins Lupus Cohort patients included in the study

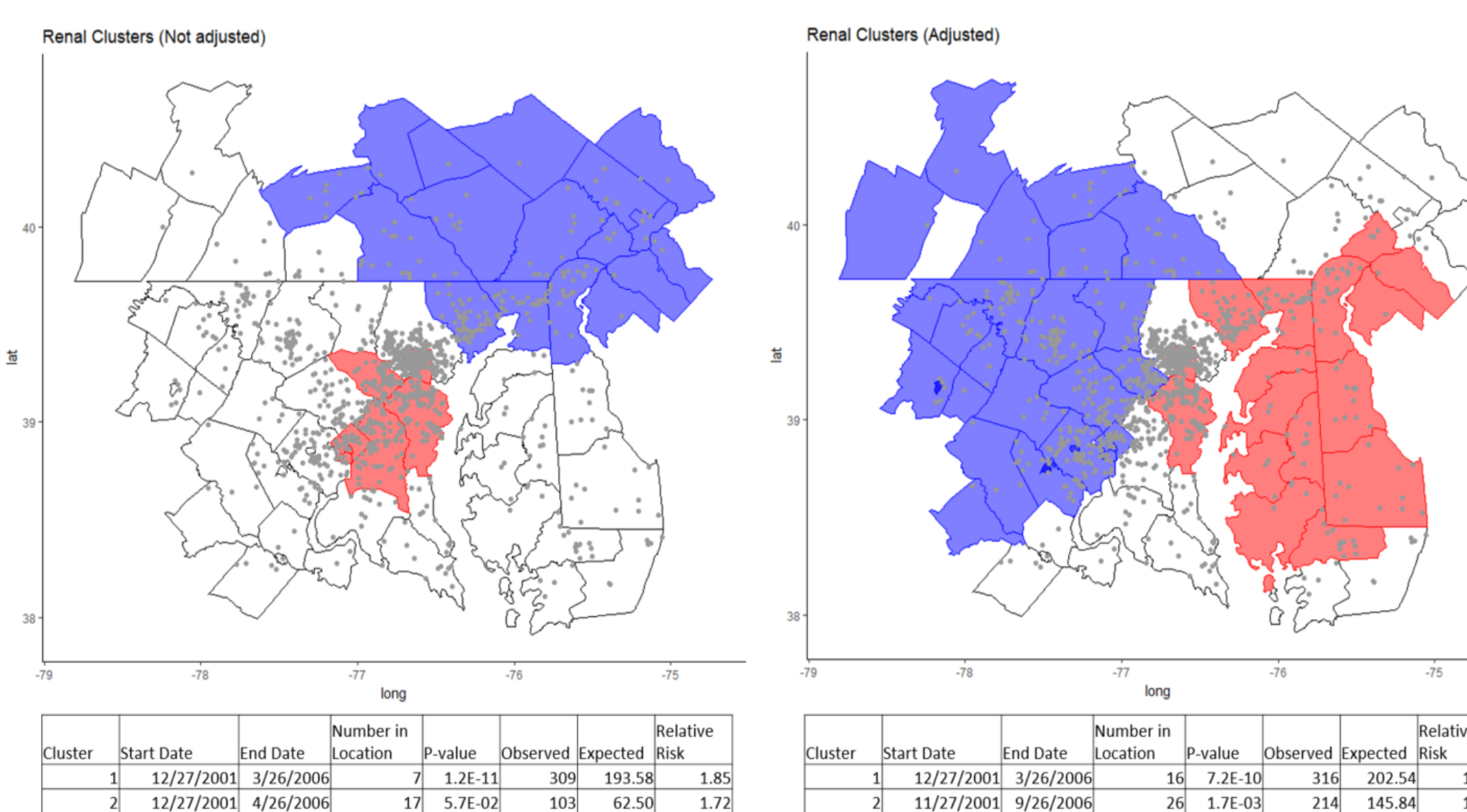


Figure 2: Renal clusters change spatially and temporally after adjusting for temperature and PM2.5

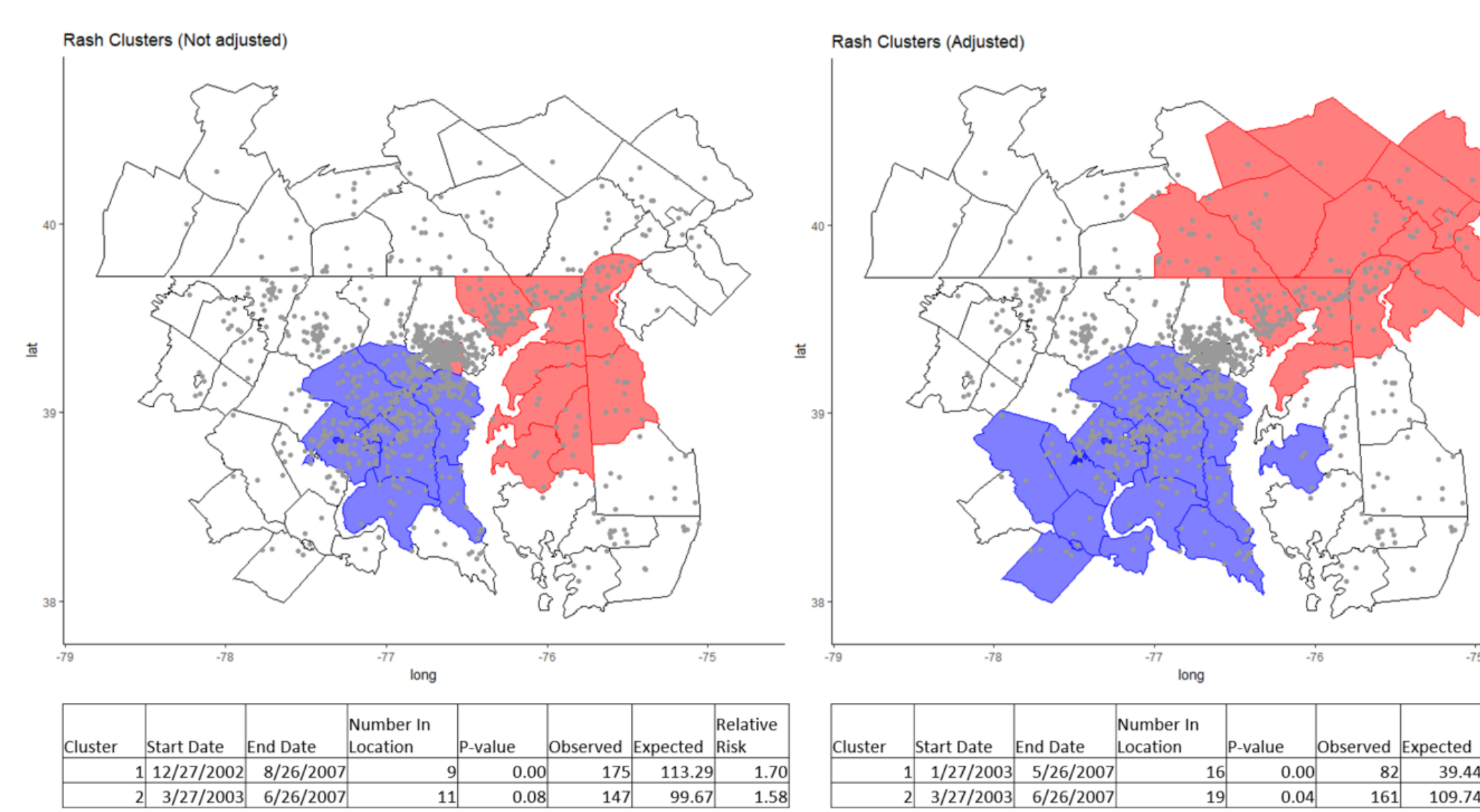


Figure 3: Rash clusters change spatially and temporally after adjusting for temperature and PM2.5

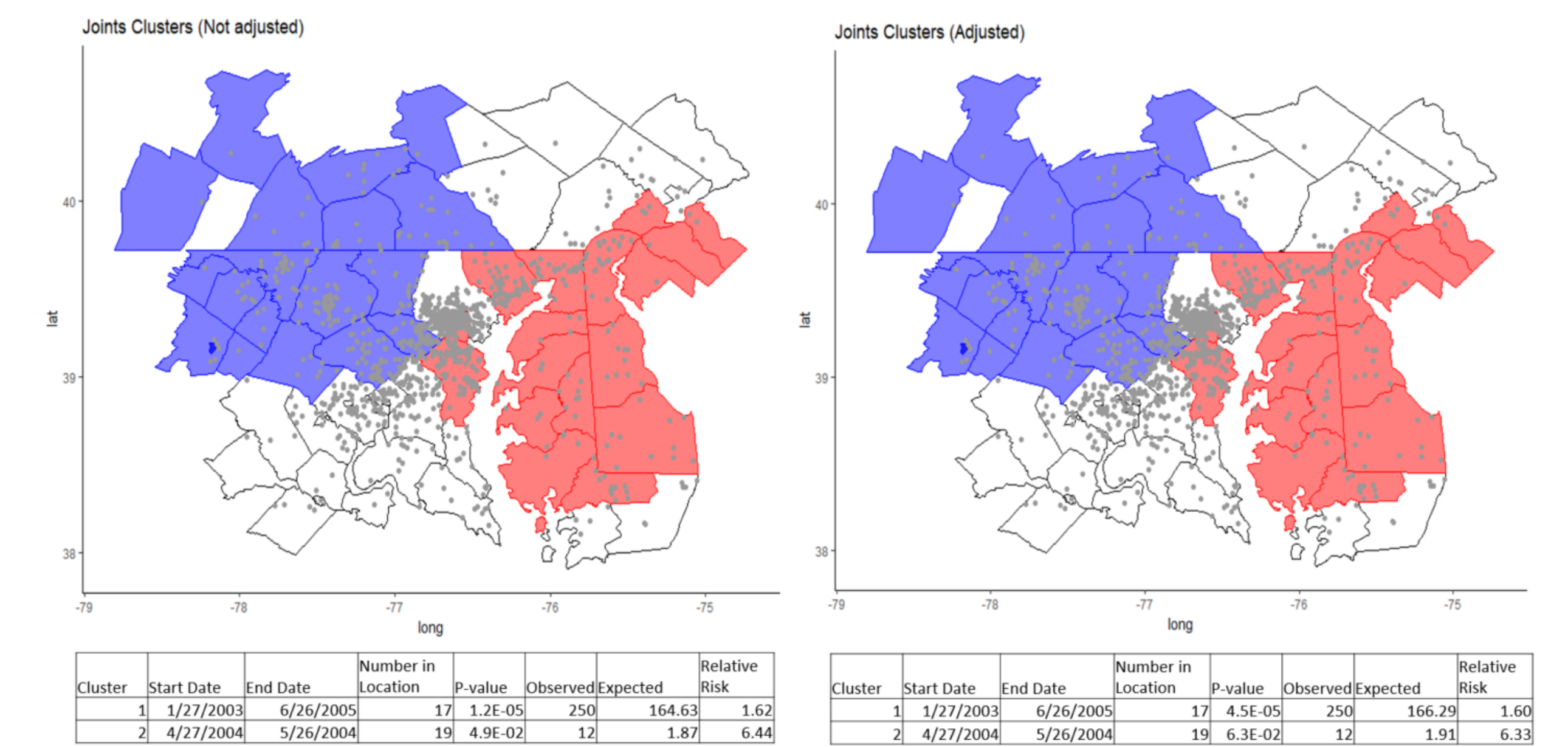


Figure 4: Joint clusters do not change spatially or temporally after adjusting for temperature and PM2.5

Rash (OR=1.0075 for 1° Fahrenheit increase), neurologic (OR=1.0096 for 1° increase), and joint (OR=1.011 for 1° increase) flares were statistically significantly associated with an increase in temperature in univariate and multivariate analysis.

Renal flares were negatively associated with increases in temperature (OR=0.996 for 1° increase) in both univariate and multivariate analysis.

Serositis flares were found to be associated in both univariate and multivariate analysis with increases in PM2.5 concentration (OR=1.024 for an increase of 1ug/m3), as were hematologic flares (OR= 1.019 for an increase of 1ug/m3), and joint flares (OR=1.011 for an increase of 1ug/m3).

After adjusting for temperature and PM2.5, rash, neurologic, and renal flare-up clusters changed spatially and temporally, suggesting that the adjustment variables could be contributing causes to the original clusters of these kinds of flare-ups.

Conclusion

An increase in temperature was found to be significantly associated with skin, joint, and neurologic flares and inversely associated with renal flares, while increase in fine particulate matter pollution was significantly associated with serositis and hematologic flares.

The clusters that remained unchanged after adjustment indicate areas of unexplained variation that requires further study.

USING LOCAL INDICATORS AND SPATIAL ANALYSIS TO ASSESS *Healthy Food Access* IN BALTIMORE CITY

BACKGROUND

Research has shown racial and socioeconomic disparities in healthy food access and subsequent risk for negative health outcomes. Policies and programs often focus on making improvements to the food environment as a way to address diet-related health disparities and reduce the number of people living in areas with low healthy food access.

The Johns Hopkins Center for a Livable Future (CLF) collaborated with the Baltimore Food Policy Initiative to develop and employ a methodology that uses primary data collection and spatial analysis to more accurately characterize the food retail environment and assess healthy food access in Baltimore City.

METHODS: *Data*

Four factors were considered for this research: distance to a supermarket*, income+, vehicle availability+, and healthy food availability across all food retail.

In order to include local level indicators and more accurately represent the food retail environment in Baltimore City, the CLF developed the Healthy Food Availability Index (HFAI) tool to measure and assess healthy foods across all food retail. The HFAI tool awards points to stores (0-28.5) based on the presence of a market basket of basic staple healthy food items.

*Baltimore City Health Department, 2016
+ United States Census Bureau, 2016 American Community Survey 5-year Estimates.

Scored items:

- Vegetables: fresh, canned, frozen
- Fruits: fresh, canned, frozen, juice
- Dried beans
- Milk: skim/low-fat, whole
- Ground beef: lean, regular
- Chicken
- Fish
- Bread: 100% whole wheat, regular
- Corn Tortillas
- Low-sugar cereal
- Rice
- Pasta
- Health frozen meals
- Low-sodium soup

METHODS: *Analysis*

A geographic area-based analysis with binary measures was used. Thresholds were determined for each factor and Geographic Information Systems (GIS) software was used to highlight areas that fell below or above each measure, indicating where residents may face physical and economic challenges in accessing healthy foods.

Each factor was combined and analyzed in grid cells in Esri's ArcGIS Desktop. The map below shows where areas fall across the spectrum of the four factors. A grid cell must meet all four factors to be categorized as a Healthy Food Priority Area. Demographic data on race and age were included to begin to identify and measure inequities.



▶ Distance to a supermarket: Areas farther than 1/4 mile from a supermarket location



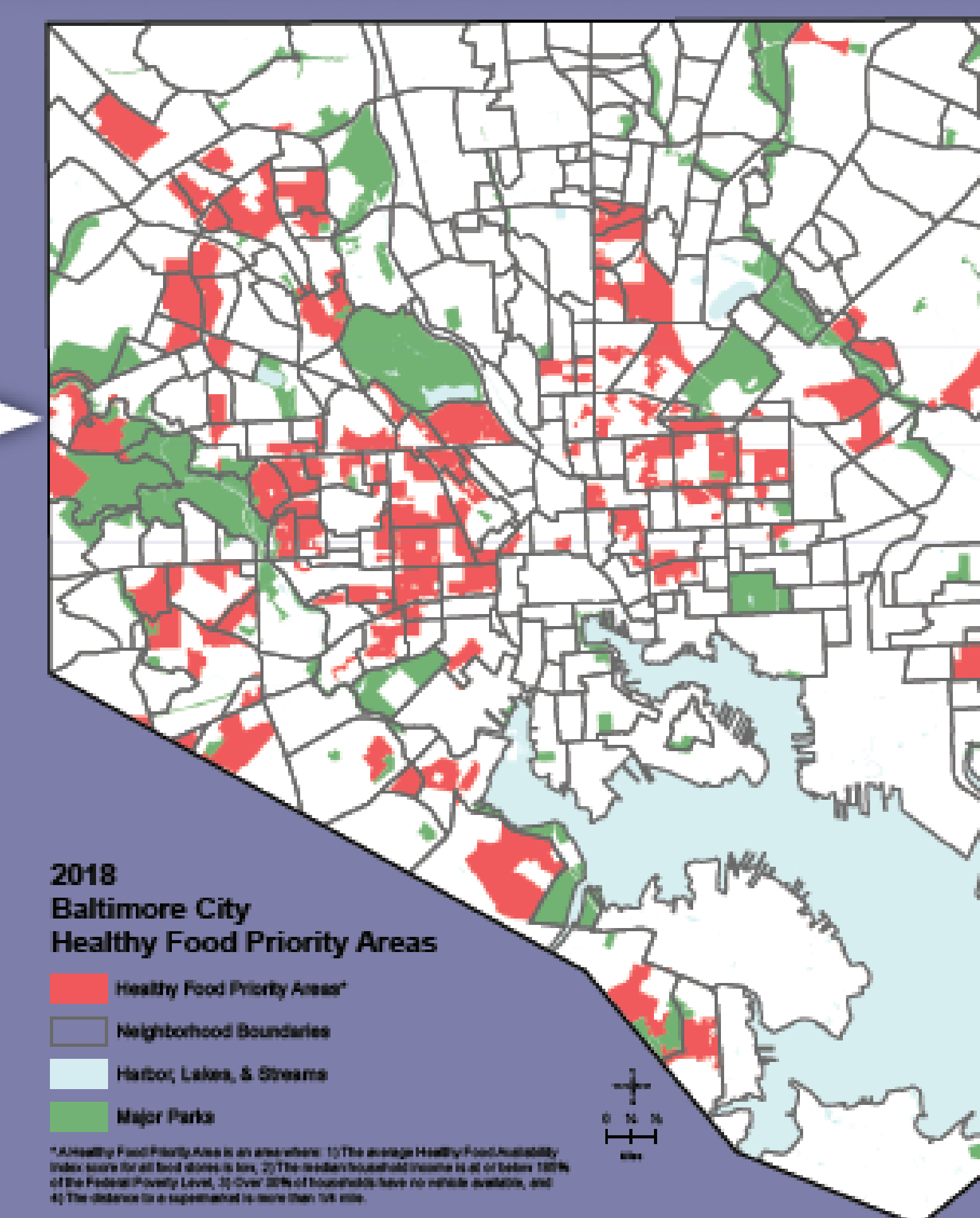
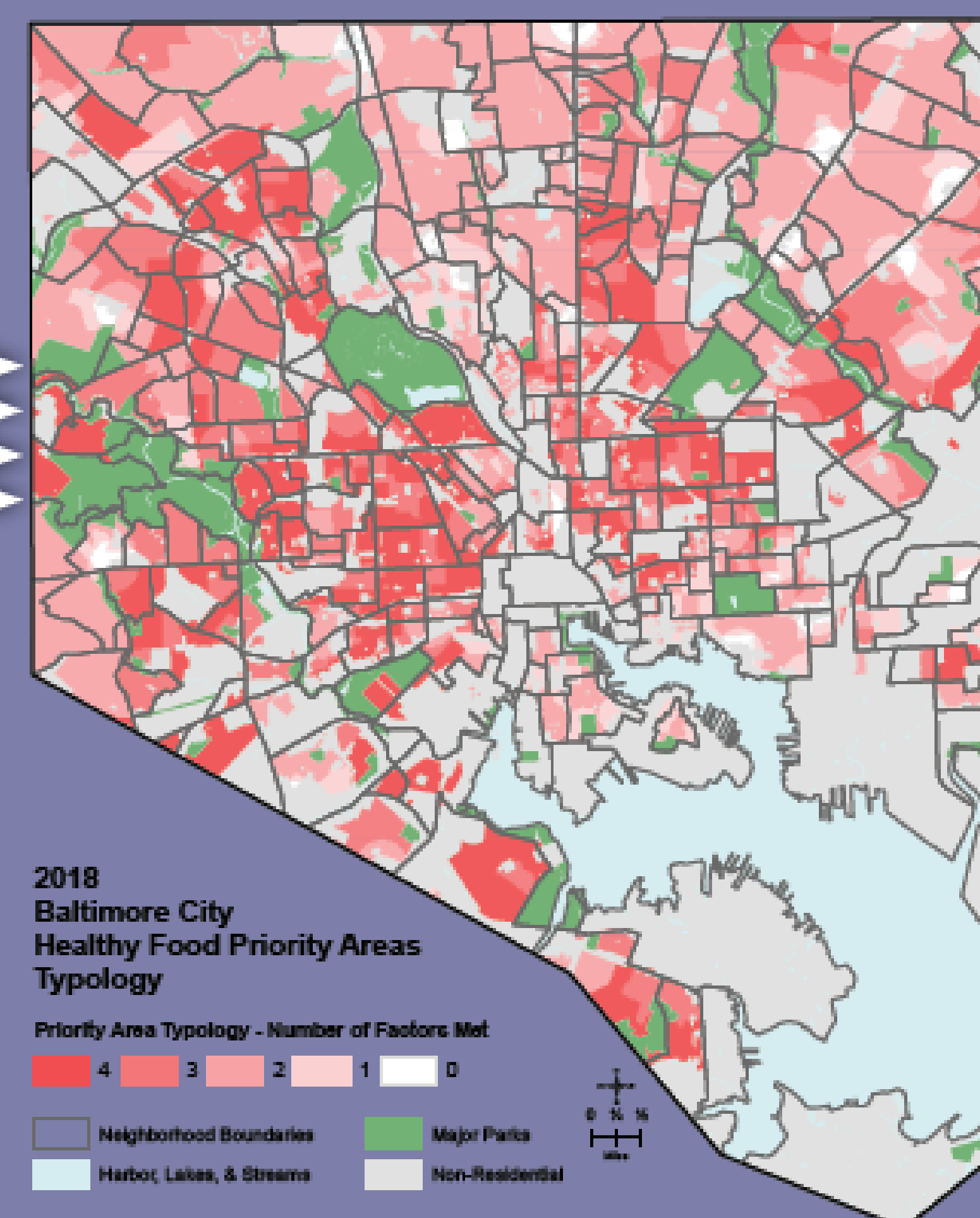
▶ Low vehicle access: 30% or more of households do not have access to a vehicle



▶ Low income: Median household income was at or below 185% of the Federal Poverty Level for a family of 4 - \$44,863



▶ Low healthy food availability in all food retail: An average HFAI score of 9.5 or below.



RESULTS

Approximately 23.5% of city residents live in areas identified as Healthy Food Priority Areas.

Black residents are the most likely of any racial or ethnic group to live in a Healthy Food Priority Area (31.5%).

Children are the most likely of any age group to live in a Priority Area (28%).

Supermarkets have the highest average HFAI score of all food retail categories and small grocery and corner stores have the widest range of scores.

CONCLUSIONS

This analysis is used to identify gaps and opportunities, inform policy and program planning, and advocate for strategies to increase healthy food access across Baltimore City. Using spatial analysis to measure food access allows government agencies, community members and decision makers to visualize geographic areas that have higher need and guide investment and energy toward appropriate policy and program solutions in these areas. This partnership between government and academia has enabled research to be translated into applied solutions and evidence-based policy change.



Using Geographical Information Systems (GIS) for Targeting Recruitment in Dementia Studies

DANNY SCERPELLA, ATIF ADAM, PHD, KATHERINE A MARX, PHD, MPH, KASEY BURKE & LAURA N GITLIN, PHD

JOHNS HOPKINS SCHOOL OF NURSING, BALTIMORE, MD
 CENTER FOR INNOVATIVE CARE IN AGING
 Funded by the National Institute on Aging grant: #NIH R01AGO41781

INTRODUCTION

An estimated 5.5 million Americans are living with Alzheimer's disease or related dementias (ADRD). Most people with ADRD are living at home, cared for by family members. As dementia prevalence will nearly double by 2050, a critical public health priority is reducing disease burden. Studies evaluating interventions are thus essential yet their quality are dependent on recruiting and enrolling eligible and representative participants. Recruitment can be challenging and costly, warranting the development and evaluation of new techniques to increase efficiencies and reduce cost. Geographical Information Systems (GIS) is one emerging approach that integrates multiple data sources and their visual representation, as well as the application of spatial analytical processes to identify potential study participants efficiently. This study evaluated the methodological efficacy of using GIS to identify neighborhoods with high density individuals at dementia risk. We layered different publically available data sources that included risk variables (e.g., gender, age, education and income).

METHODS

This study involved 234 dyads (people with dementia and family caregivers) recruited for a NIA funded behavioral intervention study in Central Maryland. Spatial analysis was performed using ArcGIS v. 10.4.1. U.S. Census data were gathered from factfinder.census.gov using American Community Survey (ACS) 2015 5-year estimates for all variables analyzed. Retrospective analysis was used to determine whether enrolled participants could be identified through GIS analyses and which variables yielded the most robust approach to identifying enrollees.

RESULTS

Figure 1

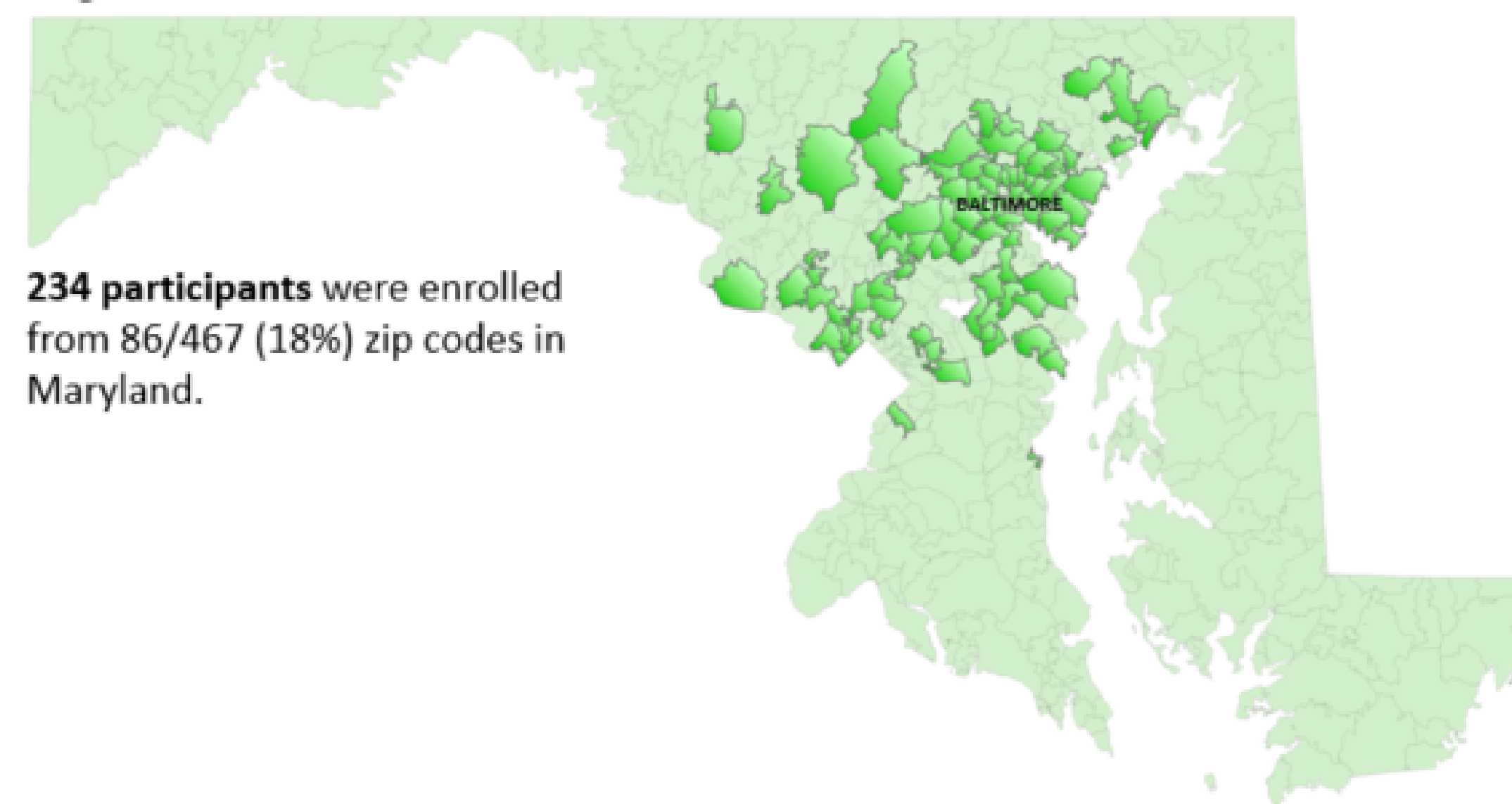


Figure 2

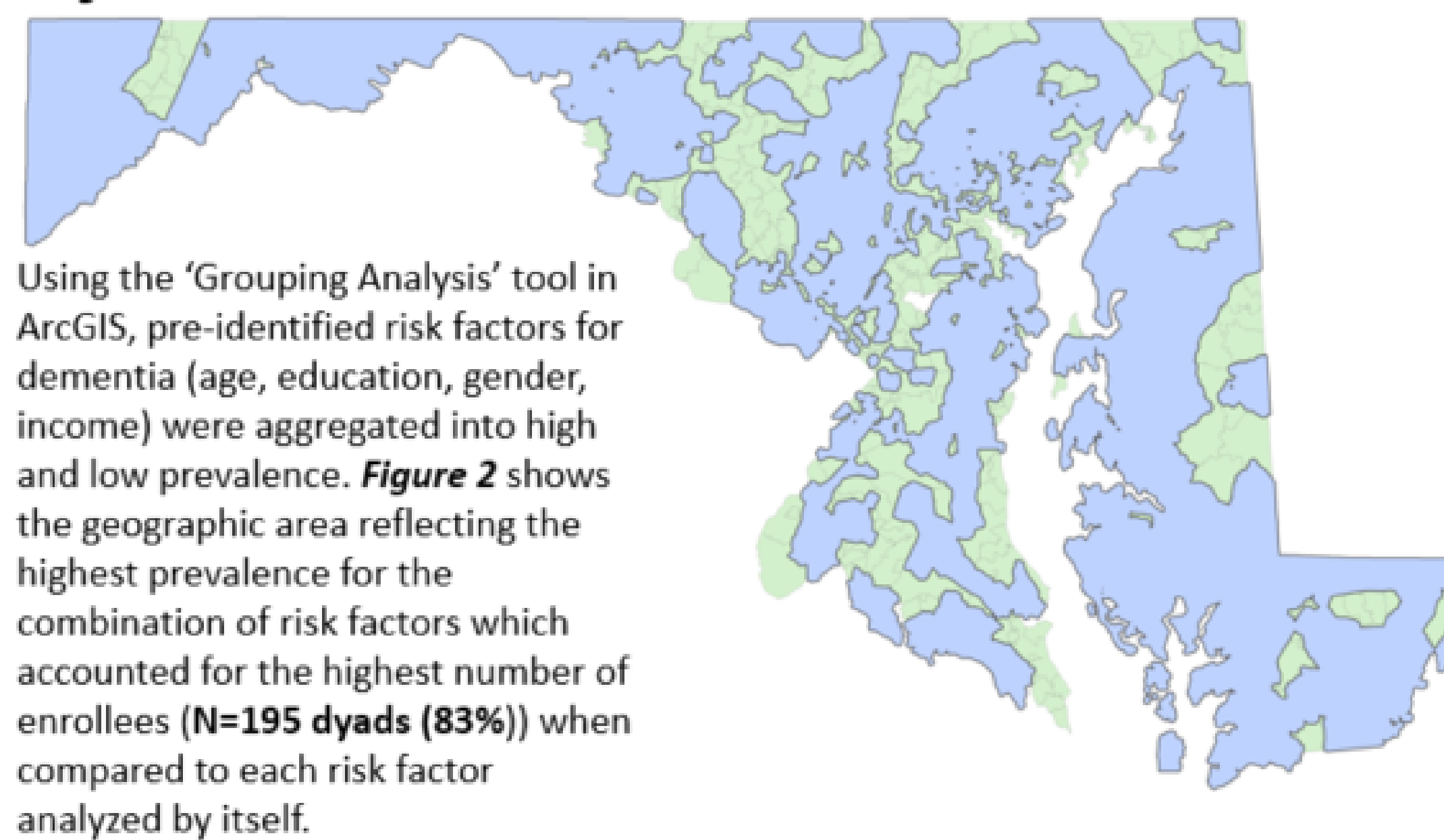


Figure 3

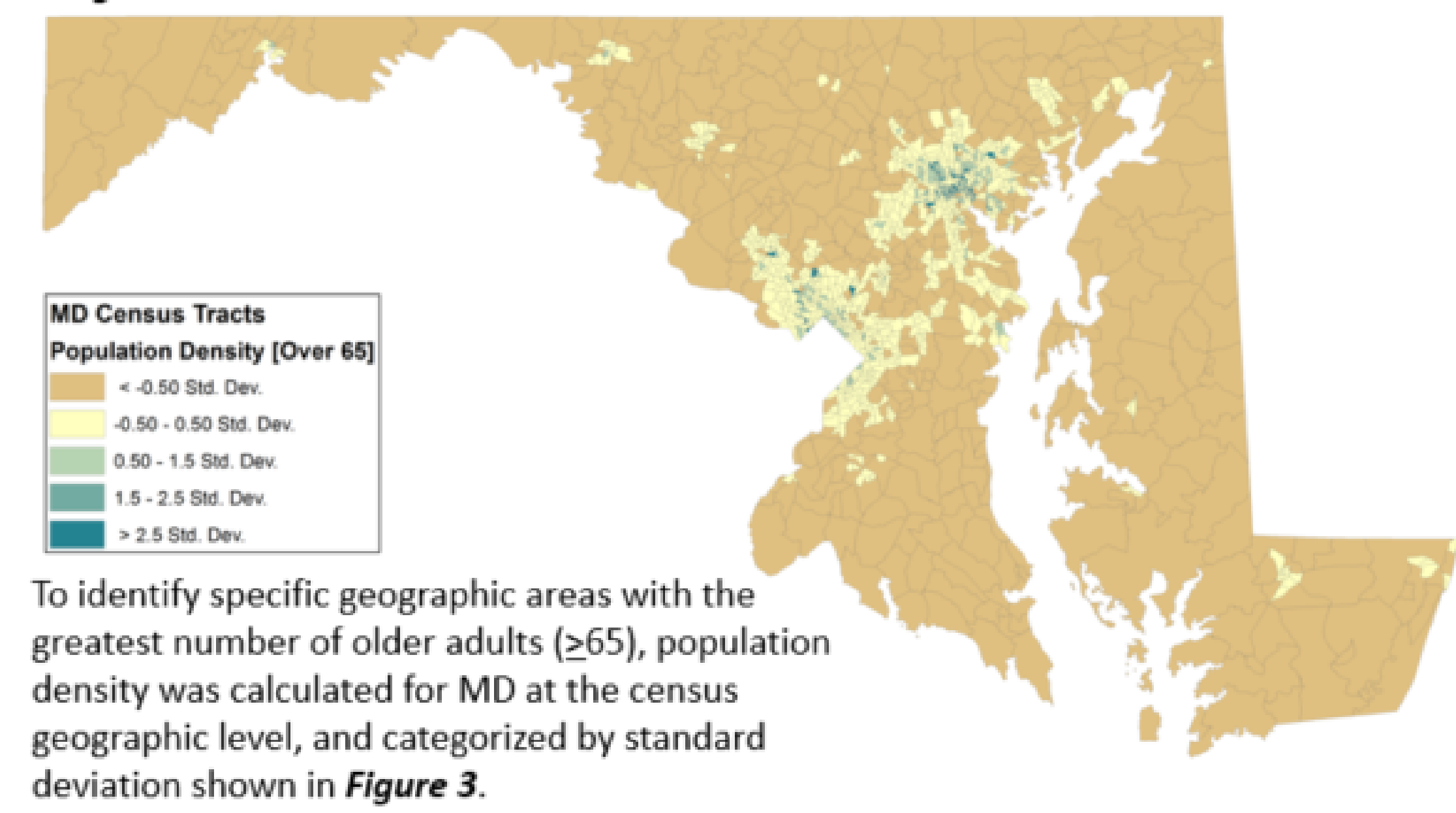
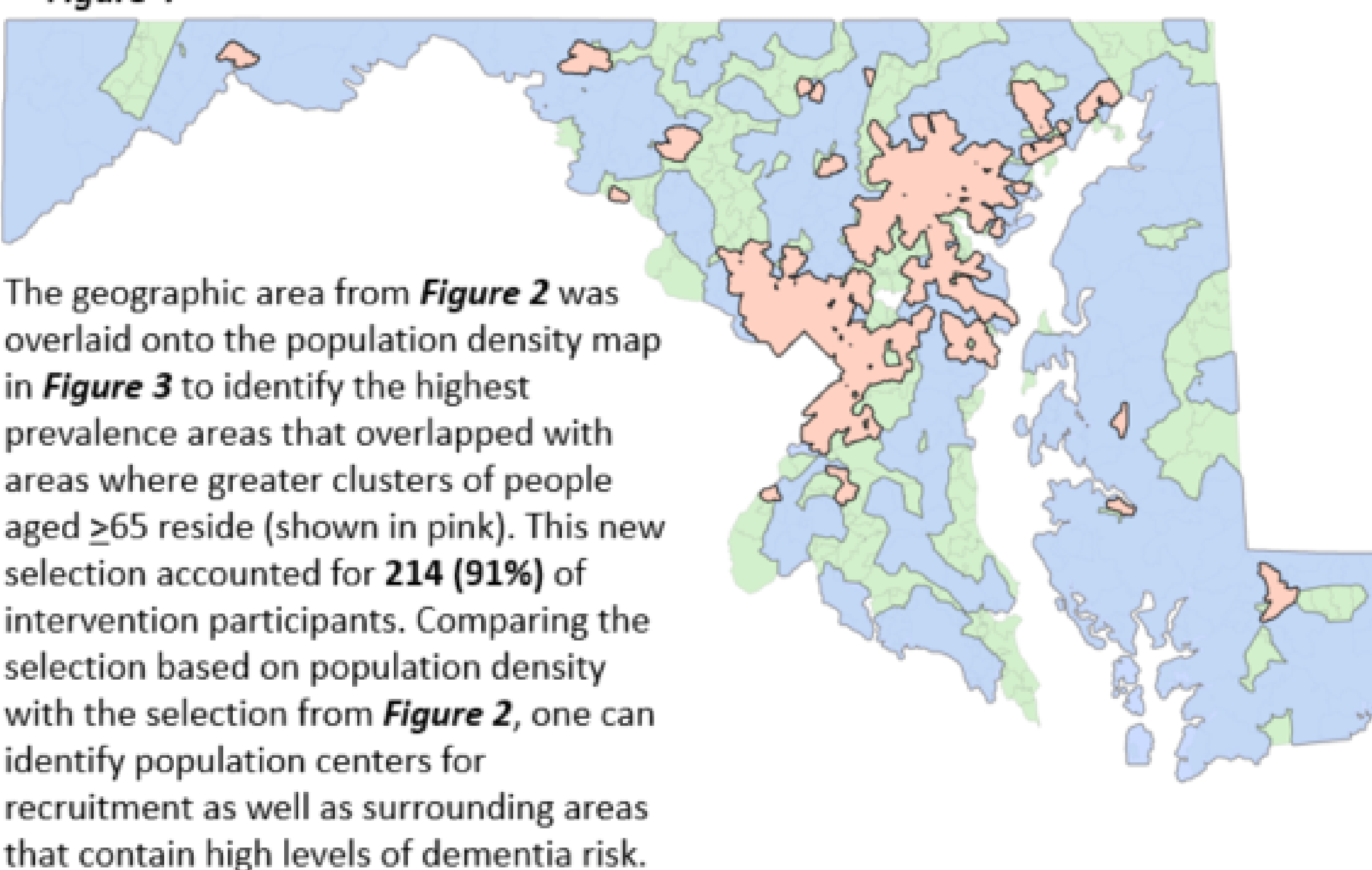


Figure 4



| | Person with Dementia | Family Caregiver |
|------------------------------|----------------------|----------------------|
| Gender (n,%) | | |
| Male | 92 (36.8%) | 47 (18.8%) |
| Female | 158 (63.2%) | 203 (81.2%) |
| Age (mean, sd, range) | 81.52 (8.00, 56-99) | 65.20 (12.64, 28-93) |
| Education (n, %) | | |
| <High School | 48 (19.2%) | 10 (4.0%) |
| High School | 67 (26.8%) | 36 (14.4%) |
| Some | 44 (17.6%) | 86 (34.4%) |
| College/Associates | 36 (14.4%) | 45 (18.0%) |
| College Degree | 54 (21.6%) | 72 (28.8%) |
| Post-graduate | 1 (0.4%) | 1 (0.4%) |
| Unknown | | |

Table 1: Demographics

CONCLUSIONS: This post-hoc analysis demonstrates that GIS was able to predict 91% of trial participants. This suggests that our methodology could be useful prospectively in designing targeted recruitment strategies. We found that identifying known risk factors of dementia such as age, gender, income, and education from available public data sets (e.g., Census data) was sufficient for analyzing and defining geographic areas where there is a high density of families living with dementia. Further evaluation of GIS is warranted as a tool to actively recruit study participants.