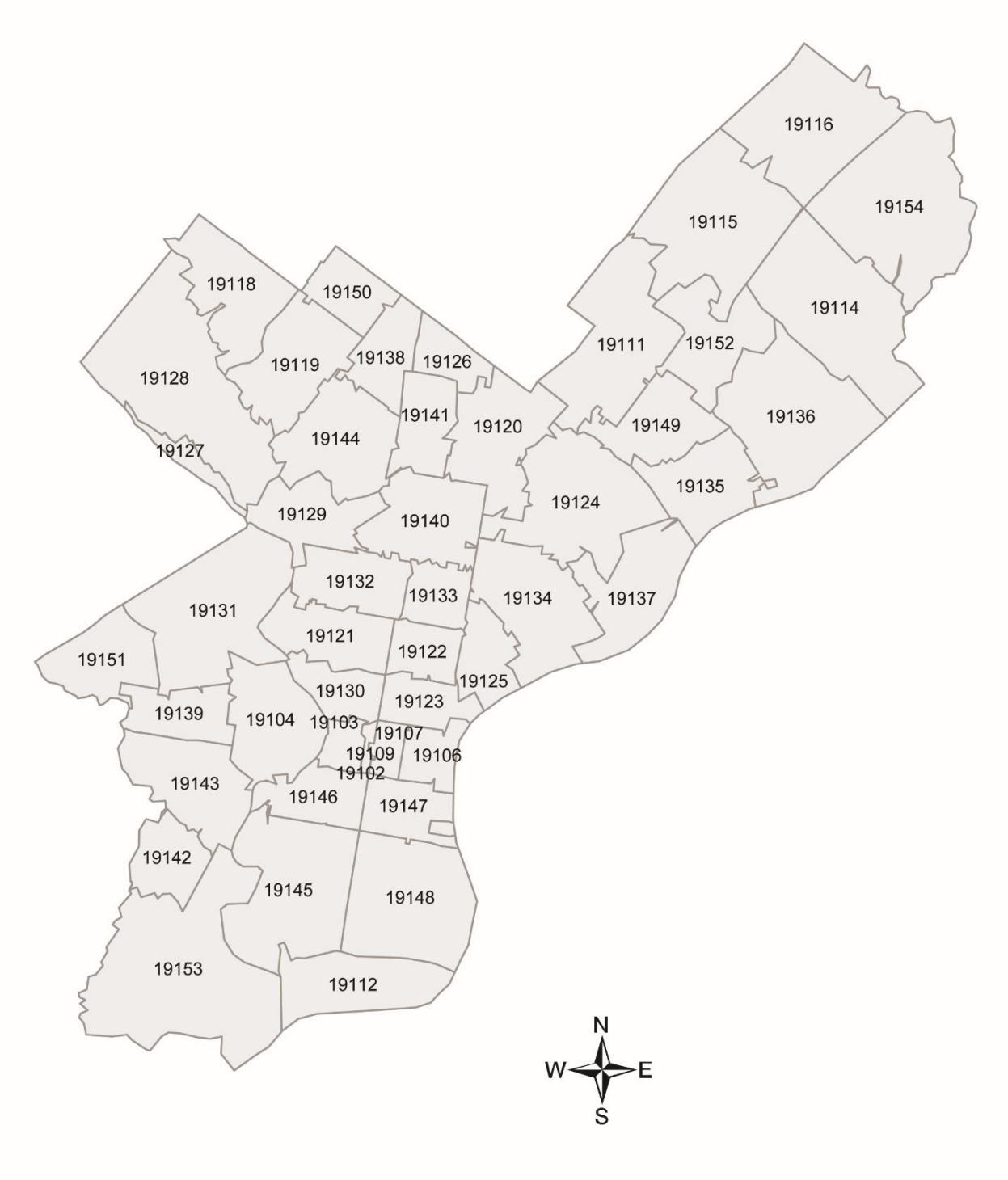
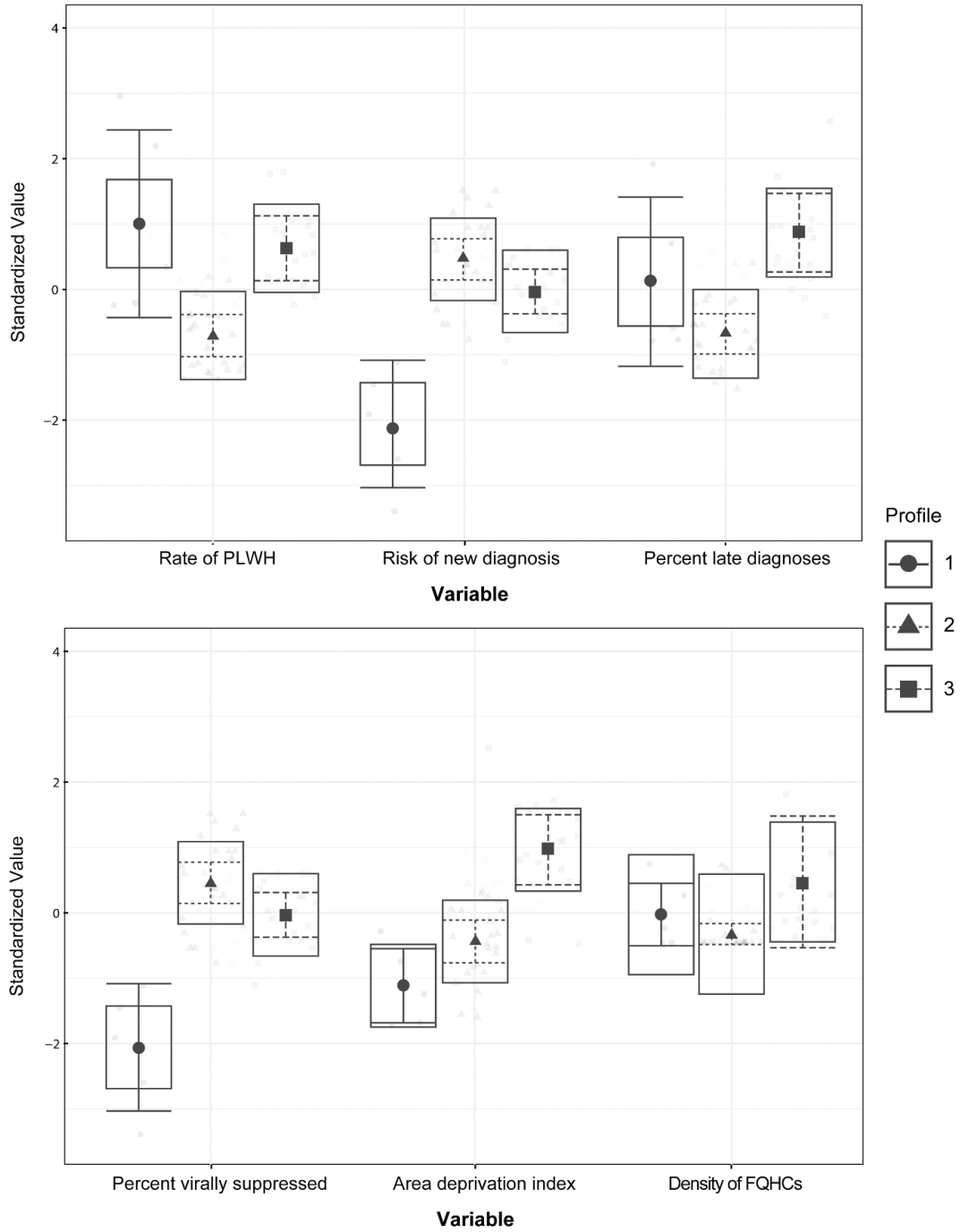
Appendix Figure 1. Philadelphia county ZIP codes reference map.



Appendix Figure 2. Latent profile analysis depicting 3 ZIP code profiles: profile 1 includes ZIP codes with residents with a high HIV treatment demand and low HIV testing demand; profile 2 includes ZIP codes with residents with low HIV treatment demand and high HIV testing demand; and profile 3 includes ZIP codes with residents with similar HIV treatment and testing demand.



FQHCs: Federally qualified health center; PLWH: Persons living with HIV.

Appendix Document 1. Additional Methodological Details on the Latent Profile Analysis and the ArcGIS Pro Location-Allocation Analysis.

Sensitivity analysis using a composite demand variable
Manifest variables used in the latent profile analysis included rates of PLWH, risk of persons newly diagnosed with HIV, percent of persons diagnosed with late HIV, percent of persons virally suppressed, density of existing health care facilities, and an area deprivation index. The area deprivation index is a standardized measure comprising ZIP code level indicators on education, poverty, female-headed households with children under 18, occupation, income, household crowding, use of public assistance or food stamps, and unemployment.[1] Population density was also considered in the latent profile analysis but subsequently removed due to poor model fit and high uniqueness in an exploratory factor analysis. The number of profiles was selected by evaluating measures of entropy and Akaike's information criterion, as well as interpretability of profiles.

The latent model of demand included three profiles (Appendix Figure 2), which may be interpreted as follows: the first represents higher treatment demand and lower testing demand (above average prevalence of PLWH, average incidence of new diagnoses, below average percentage of late diagnoses, below average percentage virally suppressed, below average area deprivation, and average density of existing healthcare facilities); the second profile represents lower treatment demand and higher testing demand (below average prevalence of PLWH, below average incidence of new diagnoses, above average percentage of late diagnoses, above average percentage virally suppressed, below average area deprivation, and below average density of existing healthcare facilities); the third profile represents similar patterns of testing and treatment demand (above average prevalence of PLWH, above average incidence of new diagnoses, average percentage of late diagnoses, average percentage virally suppressed, above average area deprivation, and average density of existing facilities).

Further details and limitations of the location-allocation analysis
The Minimize Weighted Impedance (P-Median) method was used for this analysis, which selects facility locations with the goal of minimizing the distance between the demand points and the facilities, considering any weights associated with each demand point. The methods developed by the ArcGIS Pro Network Analyst extension to solve location-allocation problems utilize heuristic as opposed to optimal procedures.[2] GIS programs have developed heuristic procedures to save computing power and solve location-allocation problems quickly.[3] It is important to note that heuristic procedures are approximate approaches to the solution, and therefore the results may be suboptimal.[3] This application demonstrated how public health professionals could apply this method to their own location-allocation problems. However, these caveats should be considered when interpreting results using heuristic procedures. While the goal of these analyses was to identify improved locations for facilities in Philadelphia, improvement to a location is subjective and should be evaluated on a case-by-case basis.

We recognized that ArcGIS may not be available to all researchers and practitioners due to its cost. Open-source tools are available to solve location-allocation problems,[4] although we have not specifically evaluated them in this work.

Workflow for ArcGIS location-allocation analysis

Data sources	
1. City of Philadelphia HIV Testing Center Locations 2. Epidemiologic dataset (ZIP code aggregated) Variables used: • Percent late diagnoses • Calculated number of demand points (estimated number of individuals living with HIV unaware of their status) ○ Sum all late diagnoses across ZIP codes, determine percentage of all late diagnoses for each ZIP code, multiply that percentage by the total number of people living with HIV unaware of their status to determine the number of demand points to allocate for each ZIP code	
ArcGIS Process	
Data import and cleaning	1. Create new project 2. Select World Street Map as base map 3. Import Philadelphia County boundary 4. Import ZIP Code boundaries a. Subset to only within Philadelphia County 5. Import City of Philadelphia HIV Testing Center Locations a. Geocode locations 6. Import Epidemiologic dataset a. Geocode ZIP code centroids 7. Create a spatial join of Philadelphia County ZIP codes and Epidemiologic data *(PhillyZIP+Epi)* 8. Create layer for bodies of water *(Water)* 9. Create new layer of Philadelphia County ZIP codes and Epidemiologic data with rivers and water removed *(PhillyZIP+Epi–Water)* a. Analysis > Tools > Analysis Tools > Erase i. Input features: *PhillyZIP+Epi* i. Erase features: *Water*
Creating population weighted demand centroids	10. Generate ZIP code centroids *(PhillyZip+Epi_Cent)* a. Import population weighted centroids i. Map > Add Data > ArcGIS Online > ZIP Code Population Weighted Centroids b. Join with *PhillyZip+Epi* data 11. Create demand weight for latent demand class a. Open attribute table for *PhillyZip+Epi_Cent* b. Calculate i. Input Table: *PhillyZip+Epi_Cent* i. Field Name: *Demand Class Weight* i. Expression Type: Arcade v. Expression: *Demand Class * Calculated number of demand points*
Allocating candidate facility locations	12. Generate evenly distributed points throughout the county, avoiding water, using Fishnet tool *(Fishnet)*

	c. Analysis > Tools > Data Management Tools > Sampling > Create Fishnet i.Template: Philadelphia County boundary i.Cell size: 1000x1000 (meters) i.Geometry type: Polygon d. Select points by location to remove points in water i.Input Features: *Fishnet_label* i.Relationship: Completely within i.Selecting Features: *PhillyZIP+Epi–Water* v.Make layer from selected features (*Fishnet_label_selection*)
Location-allocation analysis	13. Analysis > Network Analysis > Location Allocation a. Input Data i.Import facilities (candidate facilities) 1. Location Allocation tab > Import facilities 2. Add locations a. Input locations: *Fishnet_label_selection* b. Name: FID c. Facility type: candidate v.Import demand points 1. Add locations a. Input locations: *PhillyZip+Epi_Cent* b. Name: ZIP Code c. Weight: i.For percent late diagnosed analysis: *Calculated number of demand points* i.For latent demand variable: *Demand Class Weight* a. Mode: Driving time b. Direction: Towards facilities c. Cutoff: 15 (minutes)* d. Type: Minimize Weighted Impedance (P-Median) e. Facilities: 37 (number of existing facilities) f. Run

* As a sensitivity analysis, multiple cut-off values (10 minutes, 15 minutes, and 20 minutes) for the distance parameter (the maximum distance an individual would travel to a facility to receive its services) were utilized to see if this had any impact on the results of the models. Since all models proposed facilities that were on average less than 5 minutes away from demand points, changing these values did not impact the proposed locations produced by these models.

1. Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. *J Urban Health*. Nov 2006;83(6):1041-62. doi:10.1007/s11524-006-9094-x
2. ArcGIS Pro. Algorithms Used by the ArcGIS Network Analyst Extension—ArcGIS Pro | Documentation. Accessed 3/1/2022, https://pro.arcgis.com/en/pro-app/latest/help/analysis/networks/algorithms-used-by-network-analyst.htm
3. Curtin KM, Voicu G, Rice MT, Stefanidis A. A Comparative Analysis of Traveling Salesman Solutions from Geographic Information Systems. *Transactions in GIS*. 2014;18(2):286–301. doi:https://doi.org/10.1111/tgis.12045
4. Gandhi U. Locating Nearest Facility with Origin-Destination Matrix (QGIS3). https://www.qgistutorials.com/en/docs/3/origin_destination_matrix.html