DART
PHACT

Data Analysis and Research Translation (DART) Presentation of HIV/AIDS Current Topics (PHACT)
Report #8, April 2023

Using Heatmaps to Visualize Trends in HIV Diagnoses

Highlights

- The New York State Department of Health (NYSDOH) developed a simple methodology to monitor temporal increases in HIV diagnoses (DX), referred to as space-time clusters of transmission.

- This method uses existing surveillance data and widely available software (SAS and Excel) to create a monthly report that allows easy visualization of trends in new HIV DX across the state and can be easily adapted for other public health monitoring efforts.

Introduction

Detecting and responding rapidly to growing HIV clusters and preventing new HIV infections is a core function of public health entities and a key component of the Ending the HIV Epidemic (EHE) initiative.[1] The New York State Department of Health (NYSDOH) sought to develop a monthly report that would track trends in new HIV diagnoses (DX) in New York State (NYS). Temporal increases in HIV DX are often referred to as space-time clusters. A rapid public health response to concerning clusters is vital to interrupt further forward transmission. Previous methodologies relied on data from partner services, medical and community-based providers, NYSDOH data products,[2] including periodic internal reports and alerts from the Centers for Disease Control and Prevention (CDC). This new method was developed to quickly visualize areas of concern for easy and intuitive prioritization of clusters for further investigation.

Methods

Information about individuals diagnosed with HIV are entered into the CDC-funded surveillance system called eHARS (electronic HIV/AIDS Reporting System). Throughout the year, various data products and reports are generated using eHARS data.[2] The method NYSDOH developed for the monthly space-time report uses SAS code (SAS Institute, Cary, NC) to generate tables of new HIV DX by county in specific time periods stratified by various demographic characteristics. These data tables are imported into an Excel spreadsheet and displayed in different tabs using conditional formatting.

In order to create a visual display that is useful, both the area (spatial component) and the time frame (temporal component) have to be selected based on the available data on disease incidence. Since county-level data are typically used to assign surveillance and partner services staff to cases of new HIV DX, county of diagnoses was chosen as the primary geographic division for the monthly space-time cluster report. Larger regions of the state can easily be visualized by combining county-level data. The temporal component was divided into a baseline period and a reporting period or window. NYS is a large and diverse state divided into 62 counties with a wide range of HIV disease burden. A reporting period was selected to accommodate low burden and high burden counties while also being indicative of current trends in diagnoses. Due to the stochastic nature of transmission and diagnosis, a shorter reporting period can lead to false indications of increases. After some experimentation, a five-year baseline period was chosen with a six-month reporting period. The most recent six-month window is compared to the previous four, six-month windows. Since reports are generated monthly, the reporting period becomes a rolling six-month window. However, we added a month-by-month tab as well to give a finer temporal view of the reporting period.

The timing of the report was dictated by the realities of the public health surveillance process. After a positive HIV test result is reported, demographic information on the individual is collected including age, residence, risk, race, etc. As all positive HIV testing algorithm results are reportable, individuals that appear to be newly diagnosed may in fact be prevalent cases from another state or jurisdiction. Since it takes time to collect and verify this information, there will necessarily be a tradeoff between the currency of new HIV DX data and its accuracy/completeness. Typically, we wait 2-3 weeks before generating a report for a particular month – e.g., the report for January will be run in mid-February.

Custom SAS code was written to generate tables of new HIV DX from eHARS data that is stratified by the chosen time periods and county of HIV diagnosis (the current county of residence is substituted if county of DX is not yet available). A five-year total is used to calculate the average number of DX in a six-month period for each county. The current six-month period (the prior six months) is compared to this average in two ways. Both a percentage increase or decrease and the number of DX above or below the average is calculated - both measures are useful to quantify the increases and prioritize areas for investigation. When imported into Excel, conditional formatting is used to color chosen cells using the value in that cell according to a specific color scheme. We chose a red to green color mapping with yellow indicating numbers that are close to baseline. This method creates a “heat map” of the processed data and allows an intuitive and rapid visual check of the trends, see Figure 1 and Supplemental Figures 1-5.

Results

As NYS continued to see an overall decrease in the number of new HIV diagnoses, data indicated that decreases were not uniform across all geographic areas, or demographic and transmission risk groups. The original county level heat map demonstrated that overall counts of new HIV diagnoses within a region could mask increases experienced by discrete subgroup populations. Therefore, to assess potential upticks of new diagnoses by geographic, demographic and risk behavior variables, the current space-time report produced at NYSDOH includes stratification by several location and demographic variables, including: county; region – we used Ryan White Region (RWR); metropolitan/micropolitan statistical area; age at HIV diagnosis; race/ethnicity; and transmission risk category. To assess possible gaps in HIV laboratory reporting or linkage to care following diagnosis, the report is further stratified by RWR and the presence of each of the following HIV laboratory tests: viral load, CD4 count/percent, and nucleotide sequence. Jurisdictions can thus create space-time cluster reports by whatever geographic area, subgroup population, or HIV laboratory reporting requirements that are appropriate for their region and circumstances, see Supplemental Figures 6-8.

Implications for Policy & Practice

Monitoring disease transmission and incidence is a key public health function. There are many options available to discover, visualize, and prioritize increases in disease transmission including public domain and commercial products. The advantages of the method presented in this paper are flexibility, low cost, and ease of implementation.

Discussion and Conclusion

By changing the spatial and/or the temporal parameters, this simple method can be adapted for monitoring different types of public health surveillance data. For example, jurisdictions that are responsible for a city might choose zip codes or census tracts as their spatial component. Diseases with a higher rate of incidence may select a smaller temporal window such as monthly or even weekly – the flu or COVID transmission would fall into such a category.

As currently implemented, this method doesn't employ any kind of statistical test. In practice we found that statistical significance was not a reliable indicator of real increases or decreases in HIV transmission (this is especially true in low burden areas) or even of public health concern. For example, the DX of several individuals with HIV in one demographic/risk category may be of greater concern than the same number in a different category. However, statistical tests may be applied as needed.

Since this method highlights both increases and decreases from a baseline period it can be used to detect gaps in reporting and other anomalies in addition to upticks or outbreaks of disease transmission. A large decrease in DX may be due to a decrease in transmission or a provider that stopped testing in a certain area – both are potential public health concerns.

References

1. Ending the HIV Epidemic: A Plan for America https://www.hiv.gov/blog/ending-hiv-epidemic-plan-america
2. NYSDOH, HIV/AIDS Statistics in New York State https://www.health.ny.gov/diseases/aids/general/statistics/index.htm

Tables and Figures

Figure 1. Representative heat map table - numbers of new DX in each of the last five, six-month windows are on the left side of the table. On the right side, deviations from the calculated five-year mean are shown colored using conditional formatting. Orange/red cells indicate increased DX numbers above the mean. Green/light green cells indicate DX below the mean and yellow cells indicate DX close to the mean. In this example, Cattaraugus and Dutchess Counties are showing current increases (circled) that likely need follow-up investigation. Erie County is showing a prolonged decrease that might also warrant investigation. Cortland County is showing exactly the expected number of new DX over the last five six-month windows.

November 2019 NY Space-Time HIV Diagnoses Excluding NYC

County	Number of new DX in previous 6 month time windows							N increase or decrease in HIV DX for each period compared to the mean for the five previous years					
	5/17 - 10/17	11/17 - 4/18	5/18 - 10/18	11/18 - 4/19	5/19 - 10/19	5 Yr Total	5 Yr Mean	5/17 - 10/17	11/17 - 4/18	5/18 - 10/18	11/18 - 4/19	5/19 - 10/19	County
Albany	21	11	13	12	12	136	13.6	7.4	-2.6	-0.6	-1.6	-1.6	Albany
Allegany	0	1	0	1	0	4	0.4	-0.4	0.6	-0.4	0.6	-0.4	Allegany
Broome	7	11	2	3	3	60	6	1	5	-4	-3	-3	Broome
Cattaraugus	1	0	0	0	5	9	0.9	0.1	-0.9	-0.9	-0.9	4.1	Cattaraugus
Cayuga	0	0	0	2	2	12	1.2	-1.2	-1.2	-1.2	0.8	0.8	Cayuga
Chautauqua	6	2	1	5	3	30	3	3	-1	-2	2	0	Chautauqua
Chemung	2	1	1	2	0	18	1.8	0.2	-0.8	-0.8	0.2	-1.8	Chemung
Chenango	0	0	0	0	0	3	0.3	-0.3	-0.3	-0.3	-0.3	-0.3	Chenango
Clinton	1	3	1	2	2	16	1.6	-0.6	1.4	-0.6	0.4	0.4	Clinton
Columbia	2	1	1	1	1	19	1.9	0.1	-0.9	-0.9	-0.9	-0.9	Columbia
Cortland	1	1	1	1	1	10	1	0	0	0	0	0	Cortland
Delaware	1	1	0	0	1	7	0.7	0.3	0.3	-0.7	-0.7	0.3	Delaware
Dutchess	9	9	7	5	16	89	8.9	0.1	0.1	-1.9	-3.9	7.1	Dutchess
Erie	41	36	39	31	28	399	39.9	1.1	-3.9	-0.9	-8.9	-11.9	Erie

Supplemental Digital Content

Heatmap Supp Figures.pdf

New York State Department of Health
AIDS Institute
Division of Epidemiology, Evaluation, and Partner Services
Data Analysis and Research Translation
BHAE@health.ny.gov