Sharing Surveillance Data across Jurisdictions: The DC/MD/VA Model

Anne Rhodes, PhD
Division of Disease Prevention
Virginia Department of Health
Outline

• Background
• Black Box projects
• Other jurisdictional data sharing
• Future Directions
Surveillance is the conscience of the epidemic  
- Dr. James Curran
National HIV Care Continuum

HIV Care Continuum Shows Where Improvements are Needed

In the US, 1.2 million people are living with HIV. Of those:

- **Diagnosed**: 86%
- **Engaged in Care**: 40%
- **Prescribed ART**: 37%
- **Virally Suppressed**: 30%

Sources: CDC National HIV Surveillance System and Medical Monitoring Project, 2011.

*Antiretroviral therapy*

DATA and PROGRAM
Data Quality: What, Why, How?

• Surveillance, Ryan White, and other HIV data are not just utilized for funding formulas and static reports.

• Real-time tracking of diagnosis, linkage, care engagement, medication adherence and viral suppression are needed.

• Current data systems - set up artificially with barriers based on funding streams, jurisdictions, disease status, etc.
Data Accuracy/Definition

How do people get included in/excluded from Continuum of Care analyses?

• Death
• Proof of out of jurisdiction address
• No care in xx period of time?
• Modeling methods?
• Only care in xx period of time?

24% of current living cases in VA HIV Surveillance system - no lab in last 5 years (n=6,005)
Data Completeness

• Markers for care cannot all be tracked in current HIV Surveillance system

• Systems outside of health department purview often have data on care status for PLWH

• Electronic medical records/health information exchanges/all payer claims databases often available in jurisdictions
Timeliness

- NHAS - 4th Goal calls to “strengthen the timely availability and use of data”

- National viral suppression rates for 2013 for persons diagnosed with HIV as of 12/31/2012 (and alive as of 12/31/2013) released in July 2016

- AIDS.GOV site has care continuum with 2011 data
Black Box: Real Time HIV Surveillance Data

• Pilot project from Georgetown, funded by NIH

• Utilizes privacy technology for sharing surveillance data among jurisdictions where an algorithm for matching is programmed in the “black box” and returns matches of varying strengths (Exact to Very Low) to each jurisdiction
DC Metropolitan Area

Chart 1. Geographic designations for the 22 counties in the Washington metropolitan area
Black Box: Implementation

• 1\textsuperscript{st} Effort: Involved DC, MD, and VA Departments of Health (2015)

• 2\textsuperscript{nd} Effort: Involved 8 jurisdictions: DC, MD, VA, NYS, NYC, WV, DE, NC, FL (2017)

• Current efforts:
  • 18-1805 Award from CDC to Georgetown to conduct matches across all participating jurisdictions
  • DC/LA/MD/NY/VA - working on project to improve utility of matching and expand to other data
Initial Pilot: DC/MD/VA Matching


- Total (N=161,343)
- District of Columbia (N=49,326)
- Maryland (N=66,200)
- Virginia (N=45,817)

Variables:

- Last name of People Living With HIV (PLWH);
- First name of PLWH;
- Date of birth of PLWH;
- Social Security Number of PLWH;
- Hierarchical race/ethnicity assignment for PLWH; and
- Soundex of PLWH
### Matching Algorithm

#### Overview of categories & definitions used in person-matching algorithm

<table>
<thead>
<tr>
<th>Match categories</th>
<th>Variable definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact</td>
<td>if last_name and first_name and dob and ssn and sex and race</td>
</tr>
<tr>
<td>Very high</td>
<td>if (last_name and first_name and dob and sex) or ssn</td>
</tr>
<tr>
<td>High</td>
<td>if last_name and first_name and dob and (sex or race)</td>
</tr>
<tr>
<td>Medium high</td>
<td>if last_name and first_soundex and dob and sex</td>
</tr>
<tr>
<td>Med (1st definition)</td>
<td>if last_name and dob and sex and race</td>
</tr>
<tr>
<td>Med (2nd definition)</td>
<td>if last_soundex and first_soundex and dob and (sex or race)</td>
</tr>
<tr>
<td>Medium low</td>
<td>if last_soundex and first_soundex and partial_dob and partial_ssn and (sex or race)</td>
</tr>
<tr>
<td>Low</td>
<td>if last_soundex and (partial_dob and partial_ssn) and (sex or race)</td>
</tr>
<tr>
<td>Very low</td>
<td>if last_soundex and (partial_dob or partial_ssn)</td>
</tr>
</tbody>
</table>
Example input data, logs, and results

1  F  19530222  5  .......-4806  HARS000002  LANE  LOIS
1  M  19620120  5  096-65-2836  HARS000003  PARKER  PETER
1  M  19541016  5  .......-2338  HARS000004  WAYNEBRUCE
1  M  19610209  1  049-44-5374  HARS000005  BUMSTEAD  DAGWOOD
9  M  19481208  3  087-53-3451  HARS000008  CHIN  ROBERT
1  F  19420409  3  .......-3085  HARS000010  MACERA  GENNY
2  M  19550528  3  .......-4850  HARS000011  MARGOLIS  HEINE
1  M  19491118  3  816-98-7618  HARS000013  MADDANS  THOMAS
1  F  19570203  5  070-89-9503  HARS000015  JONES  CYNTHIA
2  M  19541014  3  318-90-6218  HARS000017  MADDANS  THOMAS

2015/03/25 11:11:06  VA data loaded
2015/03/25 11:11:06  Matching DC & MD
2015/03/25 11:11:08  Matches found:  15
2015/03/25 11:11:08  Matching DC & VA
2015/03/25 11:11:09  Matches found:  16
2015/03/25 11:11:09  Matching MD & VA
2015/03/25 11:11:11  Total Matches:  31
2015/03/25 11:11:11  Report created: /Data/DC/private/DC20150325111111.txt

DC-HARS000005  : MD-HARS000005  VERY_HIGH  1 2 Last First DOB Sex
DC-HARS000028  : MD-HARS000028  VERY_LOW  9 1 Soundex(Last) Partial(DOB)
DC-HARS000640  : MD-HARS000640  MEDIUM  2 1 Last DOB Sex Race
DC-HARS001299  : MD-HARS001299  EXACT  1 1
DC-HARS001401  : MD-HARS001401  VERY_HIGH  9 2 Last First Partial(SSN) DOB Sex Race
DC-HARS001554  : MD-HARS001554  MEDIUM_HIGH  1 2 Last Soundex(First) DOB Sex
DC-HARS001840  : MD-HARS001840  MEDIUM_LOW  1 9 Soundex(Last) Soundex(First) Partial(SSN) Partial(DOB) Sex
DC-HARS004523  : MD-HARS004523  VERY_HIGH  1 1 Last First DOB Sex Race
DC-HARS003257  : MD-HARS003257  HIGH  2 2 Last First DOB Race
DC-HARS006382  : MD-HARS006382  MEDIUM  2 1 Last DOB Sex Race
DC-HARS004805  : MD-HARS004805  VERY_LOW  9 9 Soundex(Last) Soundex(First) Partial(SSN) Sex
DC-HARS007579  : MD-HARS007579  LOW  2 1 Soundex(Last) Partial(SSN) Partial(DOB) Race
Black Box 1.0 Results

Output of person-matching across DC, MD, and VA eHARS databases:

<table>
<thead>
<tr>
<th>Person matches across jurisdictions:</th>
<th>Exact</th>
<th>Very High</th>
<th>High</th>
<th>Medium High</th>
<th>Medium</th>
<th>Very Low</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC-MD*</td>
<td>4013</td>
<td>5907</td>
<td>53</td>
<td>268</td>
<td>645</td>
<td>482</td>
<td>11 368</td>
</tr>
<tr>
<td>MD-VA*</td>
<td>856</td>
<td>2343</td>
<td>11</td>
<td>117</td>
<td>377</td>
<td>865</td>
<td>4569</td>
</tr>
<tr>
<td>VA-DC*</td>
<td>1064</td>
<td>3340</td>
<td>15</td>
<td>149</td>
<td>438</td>
<td>529</td>
<td>5535</td>
</tr>
<tr>
<td>Total</td>
<td>5933</td>
<td>11 590</td>
<td>79</td>
<td>534</td>
<td>1460</td>
<td>1876</td>
<td>21 472</td>
</tr>
</tbody>
</table>

*Bidirectional reporting results (i.e., DC-reported MD matches were equal to MD-reported DC-matches; etc.)

Over half of matches were not known to jurisdictions
Black Box 2.0 Results for VA: August 2017 Match, Exact and High Categories

36% of matches in exact and high categories not previously known to Surveillance program
2.0 - Other considerations

• All Document Import (ADI) files were produced by Black Box, which allowed for import directly into eHARS for each match level (1 file for exact, 1 for very high, etc)

• All jurisdictions did not import same levels of matches - issue for resolving duplicates across jurisdictions in the future

• Large amount of data to process
1805 project (funded 2018-2022)

• Currently 16 jurisdictions participating for a run in November 2018

• Some changes to matching algorithm

• Will assist with Routine Interstate Duplicate Report and Cumulative Interstate Duplicate Report

• Need to address issues of different QA and other processes across jurisdictions
Regional Data Sharing - DC/MD/VA/WV

Data Subcommittee - discusses SAS code, QA processes

Quarterly Exchanges of Data - MHS, eHARS, unmatched labs

Sharing of access to STD/HIV systems

Improved Communication/Sharing of Staffing Updates

Monthly calls among all Jurisdictions
Results: So Far

Improved Accuracy of Case Numbers

- After address and vital status updates, number of PLWH living in Virginia as of 12/31/2017 decreased by 1,594 since 2014 despite 2,776 new cases in that time period

Increased Number of Care Markers for Continuum

- Persons with a care marker increased from 53% in 2014 to 66% in 2017
- Persons virally suppressed increased from 38% in 2014 to 66% in 2017
Results: Continued

• Projects across jurisdictions, including Data to Care, cluster investigations and coordination of prevention and care efforts - quarterly in-person meetings among all HIV/STD/HCV staff

• Coordination of funding across jurisdictions and Part A/B of Ryan White to ensure needed services are provided

• Learn how other jurisdictions utilize rebates, implement Early Intervention Services, etc.
Future Directions

Continued Quarterly Data Sharing

Sharing Data Systems

National Data Sharing

Sharing across diseases
Final Thoughts

• Data Improvement strategies should be part of plan for addressing disparities on HIV health outcomes

• Sharing data across jurisdictions is important for tracking disease and care for STDs and HIV

• Utilizing data for public health impact requires merging of multiple sources of information across systems, agencies, and funding streams
Acknowledgements

CDC: Benjamin Laffoon, Dr. Irene Hall, Kim Williams, Kirk Henny, Pilgrim Spikes
DC Department of Health: Michael Kharfen, Garret Lum, Rupali Doshi, Brittany
Georgetown University: Jeff Collman, Joanne Michelle Ocampo, J Smart, Raghu Pemmaraju, Auntre Hamp
HRSA: Jessica Xavier, John Hannay
Maryland Department of Health: Colin Flynn, Reshma Bhattacharjee, James Carrier
RDE Systems: Jesse Thomas, Anusha Dayananda, RDE Developer Team
Virginia Department of Health: Lauren Yerkes, Sonam Patel, Jyoti Gupta, Celestine Buyu