Measuring Computerized Criminal History Data Quality to Improve Accuracy and Completeness

Maine State Police
Matt Ruel, Director, Maine State Bureau of Identification

SEARCH Staff
Yogesh Chawla, Manager, Software and Data Engineering Program
Becki Goggins, Director, Law and Policy Program
Mike Jacobson, Information Sharing Specialist
Mark Perbix, Director, Information Sharing Program

July 13, 2021
Purpose of the Session

• Report on the work of the CCH Metrics Working Group
• Provide update on common research and operational metrics
• Discuss common CCH data model
• Demonstrate analytic tools and capabilities
CCH Metrics Working Group

• Research using CHRI
  – Recidivism
  – Redemption

• Operational and Policy metrics using CHRI
  – Arrests missing dispositions
  – Arrest and disposition reporting patterns
  – Time from arrest to disposition
• Recidivism among persons released from state prisons has been a longstanding research topic

• Recidivism among individuals who received a non-custodial sentence compared to a jail or prison sentence has not been examined as closely

• Recidivism rates based on length of time in community is another trending research topic
Today’s Focus – Operational Metrics

Filters... Arreting Agency: AUBURN PD, AUGUSTA PD, BANGOR PD, PORTLAND PD, WATERVILLE PD, WESTBROOK PD

Missing Dispositions (15 Agencies Max)

<table>
<thead>
<tr>
<th>Agency</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUBURN PD</td>
<td>2,032</td>
</tr>
<tr>
<td>AUGUSTA PD</td>
<td>2,396</td>
</tr>
<tr>
<td>BANGOR PD</td>
<td>2,454</td>
</tr>
<tr>
<td>PORTLAND PD</td>
<td>6,027</td>
</tr>
<tr>
<td>WATERVILLE PD</td>
<td>2,048</td>
</tr>
<tr>
<td>WESTBROOK PD</td>
<td>1,959</td>
</tr>
</tbody>
</table>

Percent of Missing Dispositions (15 Agencies Max)

<table>
<thead>
<tr>
<th>Agency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUBURN PD</td>
<td>17%</td>
</tr>
<tr>
<td>AUGUSTA PD</td>
<td>20%</td>
</tr>
<tr>
<td>BANGOR PD</td>
<td>13%</td>
</tr>
<tr>
<td>PORTLAND PD</td>
<td>19%</td>
</tr>
<tr>
<td>WATERVILLE PD</td>
<td>18%</td>
</tr>
<tr>
<td>WESTBROOK PD</td>
<td>17%</td>
</tr>
</tbody>
</table>
Common Data Model (CDM)

• **Purpose of the CDM**

• **Current state of the CDM**
  – Data received to date
  – Normalization of data
  – Derived elements

• **Challenges**
  – Standardizing charges
  – Aggregating dispositions
  – Standardizing sentencing data

• **State Data Profiles**
# Example Derived Elements

<table>
<thead>
<tr>
<th>Age Range at Arrest</th>
<th>Under 10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 - 17</td>
</tr>
<tr>
<td></td>
<td>18 - 19</td>
</tr>
<tr>
<td></td>
<td>20 - 24</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
</tr>
<tr>
<td></td>
<td>40-49</td>
</tr>
<tr>
<td></td>
<td>50-59</td>
</tr>
<tr>
<td></td>
<td>60-69</td>
</tr>
<tr>
<td></td>
<td>70-79</td>
</tr>
<tr>
<td></td>
<td>80 or older</td>
</tr>
</tbody>
</table>
## Example Derived Elements

<table>
<thead>
<tr>
<th>Time from arrest to disposition in days</th>
<th>How long it took to receive a disposition for this arrest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arrests without any disposition</th>
<th>True if there is NOT a court disposition of any kind for this arrest cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Some Challenges

• Charge Mapping Because the Charge Statute was Not in the Charge Table
  – Raw data:
    Charge Statute – 99 MRSA Arson
  – Transformed data:
    Charge Statute – 17-A MRSA 802
Some Challenges

- Dispositions Reported as Counts in Separate Fields
  - Raw Data
    Final Counts = 1  Convicted Counts = 1
    Acquitted Counts = 0  Nol Pro Counts = 0
  - CDM Data
    C_Charge_Disposition_Code = CONVICTED
Some Challenges

• **Sentences in Text Strings**
  
  – Raw data
  
  Sentence detail = INCARCERATED 364 DAYS ALL BUT 30 DAYS SUSPENDED

  – CDM data elements

  SCATEGORY = INCARCERATED

  SCONFINEMENT_DAYS = 30

  SSUSPENDED_DAYS = 334
State Profiles

• **Arkansas**
  – Arrest date range: 1900 – 2020
  – Number of people in CCH: 827,793
  – Number of arrests: 2,470,192

• **Maine**
  – Arrest date range: 1931 - 2020
  – Number of people in CCH: 403,390
  – Number of arrests: 1,253,613

• **New Jersey**
  – Arrest date range: 1904 - 2020
  – Number of people in CCH: 2,277,404
  – Number of arrests: 6,301,265
Analytic Tool Demonstrations

• **CCH Dashboard**
  – Easiest to use
  – Pre-defined queries
  – Filtering capabilities

• **Microsoft Power BI**
  – Supports custom queries
  – Requires licensing
  – Large array of visualizations and export/publishing capabilities

• **Piet Query Tool**
  – Supports custom queries
  – Open-source platform
  – Limited visualizations
Summary

- Map and transform data to Common Data Model (CDM)
- Work with operations team to determine valuable metrics and visualizations
- Select visualization and BI tools
- Use operational metrics to improve Criminal History quality
Questions
Thank you

**Yogesh Chawla**, Manager, Software and Data Engineering Program
[yogesh@search.org](mailto:yogesh@search.org)
(608) 438-5965

**Becki Goggins**, Director, Law and Policy
[becki.goggins@search.org](mailto:becki.goggins@search.org)
(334) 201-3001

**Michael Jacobson**, Information Sharing Specialist
[mjacobson@search.org](mailto:mjacobson@search.org)
(406) 465-8995

**Mark Perbix**, Director, Information Sharing Program
[mark@search.org](mailto:mark@search.org)
(916) 712-5918