Linkage and Enhancement of Public Health Datasets and All Payer Claims to Further Population-Level Opioid Research

Sara Hallvik, MPH
Senior Director, Analytic Services
Comagine Health
Learning Objectives

• Data sources to better understand opioid risk factors and harms.
• Procedure to create a robust opioid research database
• Methodology prepare All Payer Claims Data to link to other sources
Background

• The opioid epidemic persists
• Fewer overdoses involve prescriptions written to the patient; more are non-medical use or illicit opioids (fentanyl, heroin)
• Does someone’s home address affect their overdose risk?
  • Do household members affect overdose risk?
  • Does community/neighborhood affect overdose risk?
Background

• Population-level opioid research using administrative data is good, but often limited
  • Breadth or depth
  • Restricted to a subset of a population (e.g. single payer type)
  • Restricted to a subset of records (e.g. paid pharmacy claims)

• Our objective was to link, at an individual patient level, public health datasets with all-payer claims and census data
  • Create rich administrative dataset
  • Enable multifaceted approach to assess prescription opioid risk
Team

• Principal Investigator: Scott Weiner, MD, MPH, Brigham and Women’s Hospital
Funding

• NIH/NIDA 1-R01-DA044167-01A1
  • PAR 16-234: Accelerating the Pace of Drug Abuse Research Using Existing Data (R01)
Aims

individual

household

community
## Data Sources

<table>
<thead>
<tr>
<th>Data source</th>
<th>Description</th>
<th>Owner</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary All Payer Claims Data (APCD)</td>
<td>Medical and pharmacy claims and enrollment data for most Commercial, Medicaid, and Medicare Advantage plans in Oregon.</td>
<td>Comagine Health Oregon Data Collaborative</td>
<td>2013 - 2018</td>
</tr>
</tbody>
</table>
# Data Sources

<table>
<thead>
<tr>
<th>Data source</th>
<th>Description</th>
<th>Owner</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary All Payer Claims Data (APCD)</td>
<td>Medical and pharmacy claims and enrollment data for most Commercial, Medicaid, and Medicare Advantage plans in Oregon.</td>
<td>Comagine Health Oregon Data Collaborative</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Census – American Community Survey</td>
<td>Publicly available American Community Survey census data</td>
<td>US Census Bureau</td>
<td>5-year average, 2014 - 2018</td>
</tr>
</tbody>
</table>
# Data Sources

<table>
<thead>
<tr>
<th>Data source</th>
<th>Description</th>
<th>Owner</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary All Payer Claims Data (APCD)</td>
<td>Medical and pharmacy claims and enrollment data for most Commercial, Medicaid, and Medicare Advantage plans in Oregon.</td>
<td>Comagine Health Oregon Data Collaborative</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Census – American Community Survey</td>
<td>Publicly available American Community Survey census data</td>
<td>US Census Bureau</td>
<td>5-year average, 2014 - 2018</td>
</tr>
<tr>
<td>Death certificate data (vital statistics)</td>
<td>Cause of death and contributing cause of death for all decedents in Oregon</td>
<td>Oregon Health Authority – Vital Statistics</td>
<td>2013 - 2018</td>
</tr>
</tbody>
</table>
# Data Sources

<table>
<thead>
<tr>
<th>Data source</th>
<th>Description</th>
<th>Owner</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary All Payer Claims Data (APCD)</td>
<td>Medical and pharmacy claims and enrollment data for most Commercial, Medicaid, and Medicare Advantage plans in Oregon.</td>
<td>Comagine Health Oregon Data Collaborative</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Census – American Community Survey</td>
<td>Publicly available American Community Survey census data</td>
<td>US Census Bureau</td>
<td>5-year average, 2014 - 2018</td>
</tr>
<tr>
<td>Death certificate data (vital statistics)</td>
<td>Cause of death and contributing cause of death for all decedents in Oregon</td>
<td>Oregon Health Authority – Vital Statistics</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Hospital discharge data (HDD)</td>
<td>Diagnostic and procedure codes recorded during each inpatient hospitalization in Oregon, regardless of payer</td>
<td>Oregon Health Authority – Office of Health Analytics</td>
<td>2013 - 2018</td>
</tr>
</tbody>
</table>
# Data Sources

<table>
<thead>
<tr>
<th>Data source</th>
<th>Description</th>
<th>Owner</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary All Payer Claims Data (APCD)</td>
<td>Medical and pharmacy claims and enrollment data for most Commercial, Medicaid, and Medicare Advantage plans in Oregon.</td>
<td>Comagine Health Oregon Data Collaborative</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Census – American Community Survey</td>
<td>Publicly available American Community Survey census data</td>
<td>US Census Bureau</td>
<td>5-year average, 2014 - 2018</td>
</tr>
<tr>
<td>Death certificate data (vital statistics)</td>
<td>Cause of death and contributing cause of death for all decedents in Oregon</td>
<td>Oregon Health Authority – Vital Statistics</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Hospital discharge data (HDD)</td>
<td>Diagnostic and procedure codes recorded during each inpatient hospitalization in Oregon, regardless of payer</td>
<td>Oregon Health Authority – Office of Health Analytics</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Prescription Drug Monitoring Program (PDMP)</td>
<td>Schedule II-V controlled substances dispensed from outpatient pharmacies in Oregon, regardless of payer</td>
<td>Oregon Health Authority – Violence and Injury Prevention Program</td>
<td>2013 - 2018</td>
</tr>
</tbody>
</table>
# Data Sources

<table>
<thead>
<tr>
<th>Data source</th>
<th>Description</th>
<th>Owner</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary All Payer Claims Data (APCD)</td>
<td>Medical and pharmacy claims and enrollment data for most Commercial, Medicaid, and Medicare Advantage plans in Oregon.</td>
<td>Comagine Health Oregon Data Collaborative</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Census – American Community Survey</td>
<td>Publicly available American Community Survey census data</td>
<td>US Census Bureau</td>
<td>5-year average, 2014 - 2018</td>
</tr>
<tr>
<td>Death certificate data (vital statistics)</td>
<td>Cause of death and contributing cause of death for all decedents in Oregon</td>
<td>Oregon Health Authority – Vital Statistics</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Hospital discharge data (HDD)</td>
<td>Diagnostic and procedure codes recorded during each inpatient hospitalization in Oregon, regardless of payer</td>
<td>Oregon Health Authority – Office of Health Analytics</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Prescription Drug Monitoring Program (PDMP)</td>
<td>Schedule II-V controlled substances dispensed from outpatient pharmacies in Oregon, regardless of payer</td>
<td>Oregon Health Authority – Violence and Injury Prevention Program</td>
<td>2013 - 2018</td>
</tr>
<tr>
<td>Emergency Medical Services (EMS)</td>
<td>Contains EMS events for most EMS agencies in Oregon</td>
<td>Oregon Health Authority – Violence and Injury Prevention Program</td>
<td>2013 - 2018</td>
</tr>
</tbody>
</table>
Linkage Procedure

• *FastLink* run in R
• Probabilistic linkage using first name, last name, DOB
• Efficiently links and de-duplicates people in very large administrative datasets
Initial Population Inclusion Criteria:
- Addresses in Oregon
- Valid zip codes
- Ages 2-100 as of 1/1/2014
- Commercial, Medicare Advantage, or Medicaid coverage

Supplemental Public Health Datasets:
- Hospital Discharge
- Vital Statistics
- PDMP

Total APCD Members
N = 3,628,992

Population Universe, by year:
2013
- Members with valid addresses n=1,970,320
- Distinct households n=968,767
- Communities n=827

2014
- Members with valid addresses n=2,417,888
- Distinct households n=1,165,316
- Communities n=827

2015
- Members with valid addresses n=2,538,576
- Distinct households n=1,196,086
- Communities n=827

2016
- Members with valid addresses n=2,621,223
- Distinct households n=1,221,231
- Communities n=827

2017
- Members with valid addresses n=2,657,851
- Distinct households n=1,243,157
- Communities n=827

2018
- Members with valid addresses n=2,528,446
- Distinct households n=1,209,236
- Communities n=828
Comprehensive Opioid Risk Registry (CORR)

**STEP 1**
Vital Records
N = 206,000 Deaths
- First Name
- Last Name
- DOB
  - Death certificate data for all OR - used to identify total opioid overdoses.

**STEP 2**
Hospital Discharge
N = 2.2 M Discharges
- First Name
- Last Name
- DOB
  - All hospitalizations in OR - used to identify opioid-related hospitalizations.

**STEP 3**
Minimally Necessary APCD
- First Name
- Last Name
- DOB
  - Comagine analyst generates binary and categorical variables; removes source variables; creates minimally necessary database. DOB changed to calculated age.

**STEP 4**
PDMP
N = 3.2 M Rxs Dispensed
- First Name
- Last Name
- DOB
  - Schedule II-V substances dispensed from OR pharmacies - used to characterize Rx use.

**STEP 5**
CORR
- De-identified, minimally necessary
  - Only Comagine staff with data use rights can access and use the data

**US Census**
*Reference table
- FIPS Code
  - SES indicators for each individual’s neighborhood by census tract.

**APCD**
N = 3.6 M People
- First Name
- Last Name
- DOB
- FIPS Code
  - Outpatient, inpatient, ED, and pharmacy claims for 80% of Oregonians. Used to identify base cohort and opioid-related ED Visits

**Enhanced APCD**
- First Name
- Last Name
- DOB
  - OHA & Comagine analysts link HDD and APCD at OHA facility.

**Minimally Necessary Enhanced APCD**
- First Name
- Last Name
- DOB
  - OHA & Comagine analysts link PDMP and Minimally Necessary Enhanced APCD at OHA facility.

**CORR**
- Patient IDs
- Provider IDs
- Pharmacy IDs
  - OHA analyst ensures patient, provider, and pharmacy IDs are removed.
Defining Households

• Household grouper (Aim 2)
  • Unique patients linked with household members in 12-month periods (April-March)
  • Uses exact address, P.O. Box, apartment number, etc.
  • Create unique ID for every household in each 12-month period
Defining Communities

• Community identifier (Aim 3)
  • Code in R runs a cyclical process
    • Submits exact address to census website
    • Converts address to latitude, longitude and FIPS code
  • Resulting output is dataset with patient ID, address, latitude, longitude and FIPS code
  • FIPS code used to pull in census tract community characteristics from census data for each person in APCD cohort
Significance

• Population-level data linkage requires substantial preparation and cleaning

• Linked datasets provide valuable information
  • Prescription and clinical history across payers with other factors predictive of overdose, and best capture of overdose events

• Other states could replicate our methodology to create a state-specific CORR
Gaps

• CORR currently only includes overdoses where person was either transported to the ED/hospital or died
  • We are planning to incorporate EMS data in CORR to fill this gap – it will provide information on opioid overdoses for patients who refused transport.
Manuscripts

• An article currently under review examines ambulance trips for substance-related issues before and after COVID-19
• Below is an in-press article examining mortality after treatment with naloxone by EMS

One year mortality of patients treated with naloxone for opioid overdose by emergency medical services

Scott G. Weiner, MD, MPH, Olesya Baker, PhD, Dana Bernson, MPH & Jeremiah D. Schuur, MD, MS
Thank you!

Sara Hallvik, MPH
Comagine Health
shallvik@comagine.org