The More You Know:
Linkage of Public Health Datasets and All Payer Claims to Further Population-Level Opioid Research

Sara Hallvik, MPH[1] • Sanae El Ibrahimi, PhD, MPH[1] • Michelle Hendricks, PhD[1] • Nazanin Dameshghi, MS[1] • Carissa Bishop, MPH[1]
• Christi Hildebran, LMSW[1] • Scott G. Weiner, MD, MPH[2,3]

[1]Comagine Health   [2]Brigham and Women's Hospital   [3]Harvard Medical School

Comagine Health

BRIGHAM HEALTH
BWH BRIGHAM AND WOMEN'S HOSPITAL

Research Objective

Administrative datasets used for opioid research are often limited, restricted to a subset of a population (e.g., a single payer type) or 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 to create a richer administrative dataset that would allow for a multi-level assessment of patient, prescription, household, and community-level predictors of fatal and nonfatal opioid overdoses.

Population Studied

Adults in the voluntary Oregon All Payer Claims Dataset (APCD) housed at Comagine Health. The APCD includes outpatient, inpatient, ED, and pharmacy claims for 81% of Oregonians.

Approach

In compliance with various data governance statutes, Comagine worked closely with the Oregon Health Authority (OHA) Public Health Division to develop a linkage plan for linking the APCD to multiple state administrative datasets to create the Comprehensive Opioid Risk Registry (CORR), which spans 2013-2018.

Step 1
- Comagine probabilistically linked (using the fastLink package in R) patient-level APCD data to Oregon Vital Records data using name and date of birth. Socioeconomic status indicators for the patients' census tract from the US Census were linked using FIPS codes. The resulting dataset provides data on fatal opioid overdoses, opioid-related ED visits, community and household groupers, condition flags and comorbidity estimates, and community-level covariates.

Step 2
- State statutes limit the sharing of identifiers for Hospital Discharge Data, so a Comagine analyst visited the OHA to conduct patient-level linkages with an OHA analyst. These data identify opioid-related hospitalizations regardless of payer.

Step 3
- In preparation for linkage to the Prescription Drug Monitoring Program (PDMP) data, the Comagine analysts are currently creating a minimally necessary dataset to replace source variables with binary and categorical variables to limit the possibility of re-identification after PDMP linkage. After this is complete, OHA reference datasets will be destroyed.

Step 4
- A Comagine analyst will visit the OHA facility to link the PDMP to the minimally necessary enhanced APCD. These data will allow the study team to characterize prescription opioid use at the patient, community, and household levels.

Step 5
- Finally, all patient, provider, and pharmacy identifiers will be removed from the final CORR dataset, removing the possibility of linkage back to the APCD.

Implications

Using public health, medical claims, and publicly available datasets, other states could replicate our methodology to create a state-specific CORR. This is a significant undertaking that requires both 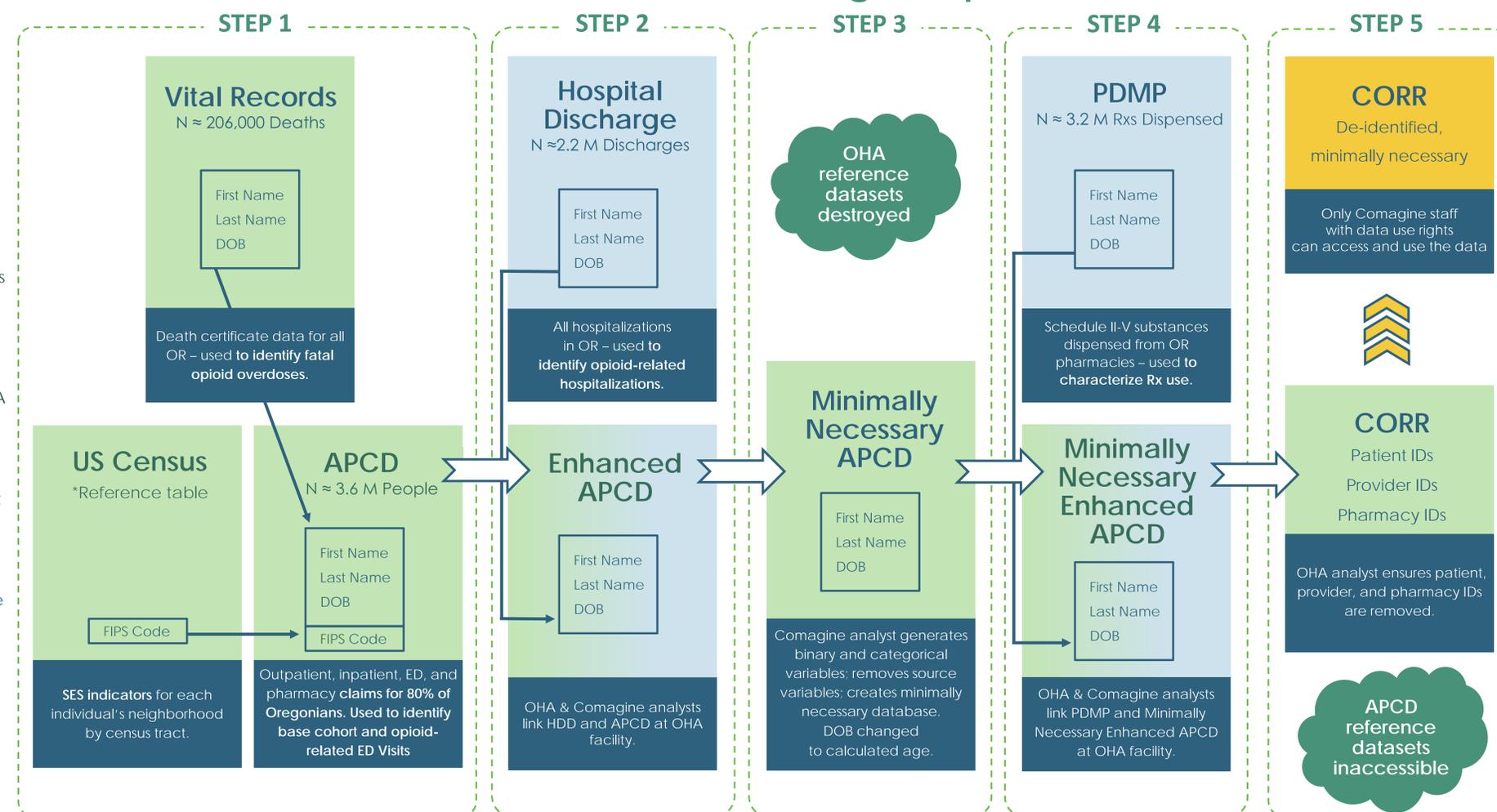financial and stakeholder support. Ultimately, these databases will be used to further opioid research with the goal of understanding predictive factors to prevent overdoses.

Principal Findings

Comprehensive opioid research datasets require a substantial amount of preparation and cleaning, but can yield valuable information. Databases like the CORR are unique in that they link prescription and clinical history across payers with other factors predictive of overdose.

CORR Data Linkage Map

STEP 1

Vital Records
N ≈ 206,000 Deaths

First Name
Last Name
DOB

Death certificate data for all OR – used to identify fatal opioid overdoses.

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

STEP 2

Hospital Discharge
N ≈ 2.2 M Discharges

First Name
Last Name
DOB

All hospitalizations in OR – used to identify opioid-related hospitalizations.

Enhanced APCD

First Name
Last Name
DOB

OHA & Comagine analysts link HDD and APCD at OHA facility.

STEP 3

OHA reference datasets destroyed

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.

Minimally Necessary Enhanced APCD

First Name
Last Name
DOB

OHA & Comagine analysts link PDMP and Minimally Necessary Enhanced APCD at OHA facility.

STEP 5

CORR
De-identified, minimally necessary

Only Comagine staff with data use rights can access and use the data

CORR
Patient IDs
Provider IDs
Pharmacy IDs

OHA analyst ensures patient, provider, and pharmacy IDs are removed.

APCD reference datasets inaccessible

Legend:
- OHA Analyst
- Comagine Analyst
- Dataset/Database Description
- Final Linked Database
- Linkage pathways and linkage variables
- Linked datasets become a single database
- Removal of all patient, provider, and pharmacy identifiers