Leveraging other data: Linking and benchmarking

Cynthia Kornegay, Ph.D.
Lead, Prescription Drug Abuse Team
Office of Surveillance & Epidemiology.
Center for Drug Evaluation and Research
US Food and Drug Administration

Introduction

- Meeting began with a discussion of data resources for evaluating ADF opioids
  - Identification
  - Sampling
  - Study Designs
  - Metrics/Measures
  - Analyses
  - Confounders
- This session will focus on ideas for leveraging those and other data resources together to enhance ADF opioid investigations

Overview

- Linking or benchmarking existing data to other data resources can serve several purposes
  - Providing additional statistical power for rare outcomes (i.e., Claims/EMR→NDI)
  - Providing additional risk factor and/or confounder information for a more comprehensive assessment (e.g., NPDS→Claims/EMR→NDI)
  - Providing additional context to address sampling and generalization issues (i.e., TEDS→N-SSATS)

Additional Statistical Power

- Large collections of healthcare information have a large number of individuals at low risk of abuse-related overdose or death
  - Individuals still participating in society
- Smaller practices that include pain and addiction management have wider variety of patients but data are less accessible

Additional Statistical Power - Example

- Study examining risk of misuse, abuse, and addiction in patients treated with extended-release and long-acting opioids for chronic pain
  - Prospective cohort study
  - Multiple complementary data resources used to
    - Enhance target population of individuals on long-term therapy for chronic pain
    - Ensure adequate participation of individuals perceived to be at high, medium, and low risk of misuse, abuse, and/or addiction

Additional Risk Factors

- Prescription drug abuse has a variety of risk factors, and outcomes may not be predictable
- Not all of these potential confounders available in health care data
  - BRFSS
  - NFLIS
  - YBRFSS
  - drug testing data
- Possible to link abuse-related data to behavioral/laboratory/law enforcement data?
- Example: FDA supporting pilot project in state of CT

Additional Risk Factors - Example

- Effort to link exposure, treatment, and outcome information statewide
  - Proof of concept – will link clinical and non-clinical data resources
  - Goal is also to determine if products prescribed to patients (specifically opioids) are the same ones involved overdoses

Additional Risk Factors
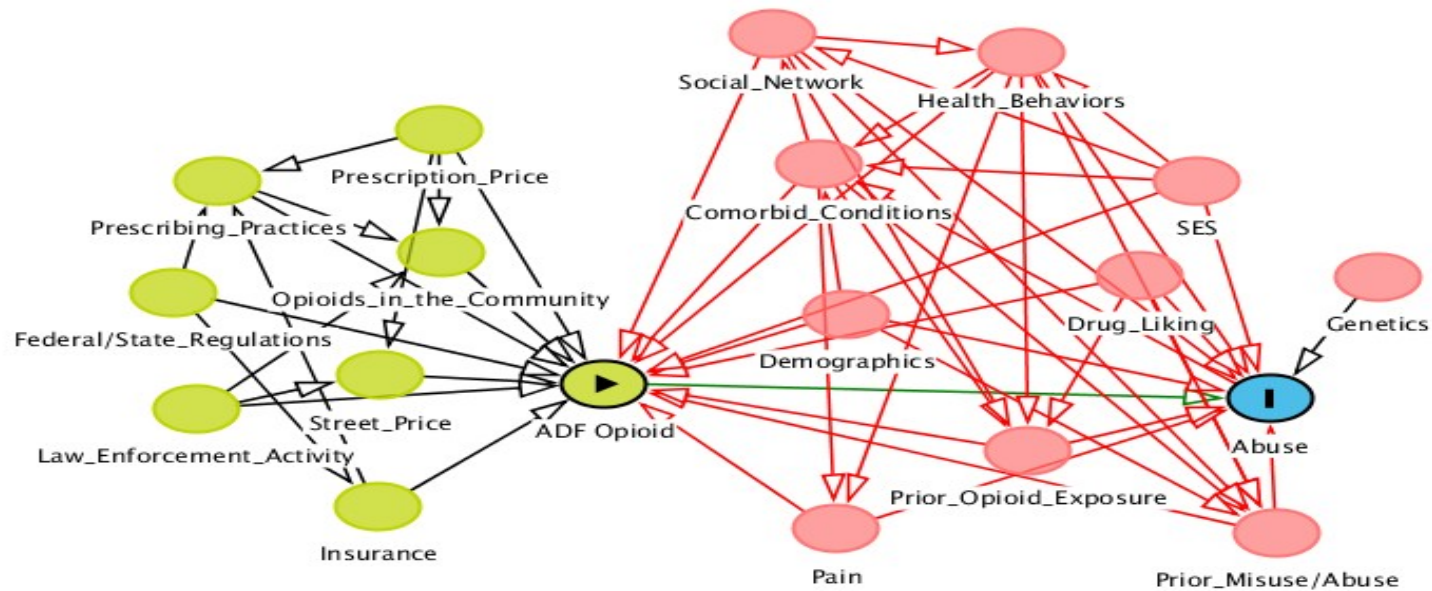
Additional Context

- Goal would be to benchmark convenience sample to enhance generalizability

- Would need to find data sources with well-defined base population of interest, and determine important variables to estimate

- Unclear that census data most appropriate
  - No drug-related data
  - Key census characteristics?

Additional Context - Example

- Outcome-based evaluation of methodology to obtain U.S.-based inpatient drug utilization levels
  - ICD9 codes and drug utilization from a large convenience sample of acute care hospitals in the US
  - Compared to National Hospital Discharge Survey
    - Nationally projected discharges
    - ICD9 codes but no drug use
  - Able to compare demographics, diagnoses, and procedures between the two samples
  - Facilitated understanding of representativeness of convenience sample
- Did not require actual linkage of data

Challenges Specific to Leveraging Data

- Even within sphere of clinical data, many and varied complications in linking data resources:
  - Technical
  - Update frequency
  - Confidentiality
  - Contractual restrictions
  - Consistency

- As a result, most linkages use data collected for similar reasons

- Linking dissimilar data will be even more difficult

Discussion Question

- Discuss potential data sources and considerations for linking to additional data sources to increase study size and power
  - Misuse and abuse definitions
  - Outcomes of interest
  - Data granularity

Discussion Question

- Discuss potential data sources and considerations for linking to additional data sources to provide additional risk factors and insights
  - Confidentiality concerns
  - Time-dependent variables

Discussion Question

- Discuss potential data sources and considerations for benchmarking to other data sources to enhance generalizability
  - Use of contextual variables (e.g., census data)