Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data

National Center for Health Statistics

Division of Health Care Statistics, Office of Analysis and Epidemiology, and Division of Vital Statistics

Technical Expert Panel Meeting

December 10, 2018

Agenda

- Overview of project goals and task
- Task 1 updates
- Task 2 updates
- Discussion

Project Goal

To improve public health surveillance and expand researchers' access to data on opioid-involved health outcomes by developing enhanced methods that make use of available structured and unstructured data from

- the National Hospital Care Survey (NHCS),
- the National Death Index (NDI), and
- the National Vital Statistics System restricted mortality data, drug specific information (NVSS-M-DO)

to identify specific opioids (e.g., fentanyl and heroin) involved in outcomes such as drug-related hospital visits and drug poisoning deaths.

Project Tasks

1. Create a merged 2014 NHCS/NDI/NVSS-M-DO file and make file available in the NHCS RDC.
2. Develop methods of enhanced opioid-identification in hospital and death certificate data.
3. Create a merged 2016 NHCS/NDI/NVSS-M-DO file with enhanced opioid identification methodology and make file available in the RDC.
4. Disseminate and promote new resources.

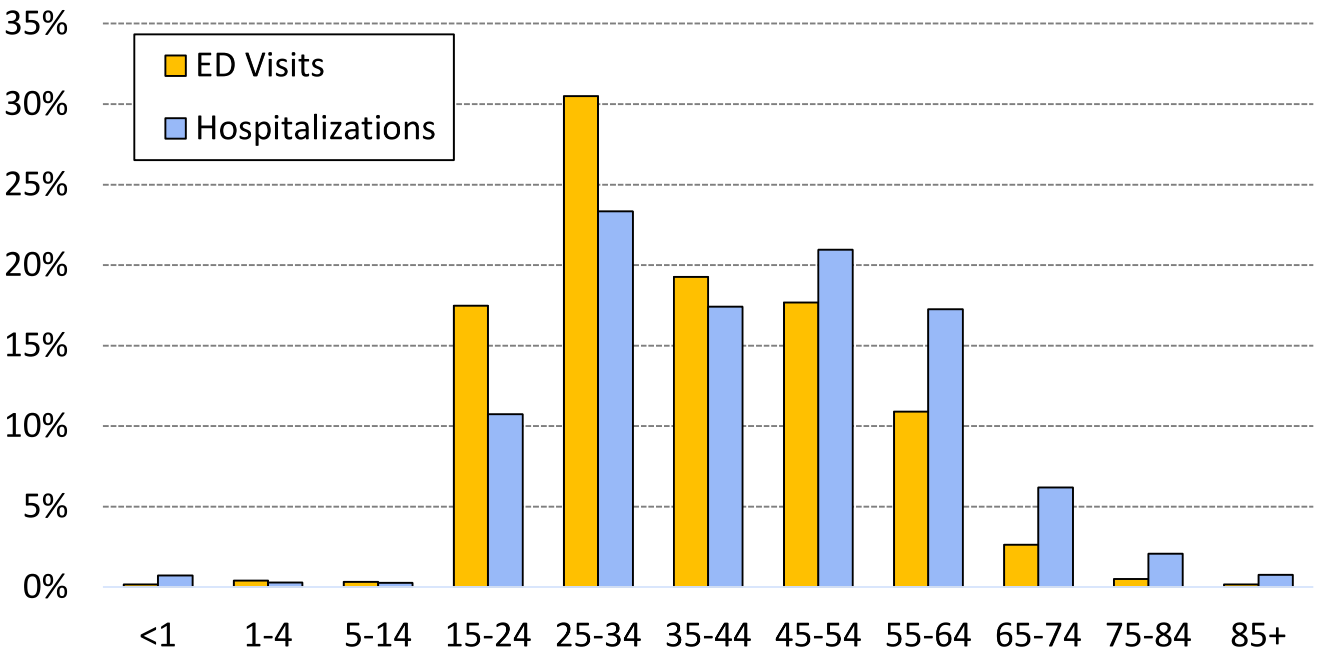
Task 1 Update

2014 NHCS/NDI/NVSS-M-DO file

Task 1 Update

- The 2014 NHCS data has been successfully linked to the 2014-2015 NDI and the 2014 NVSS-M-DO data.
- Staff are continuing to analyze the merged file in preparation of a white paper to be released March 2019.

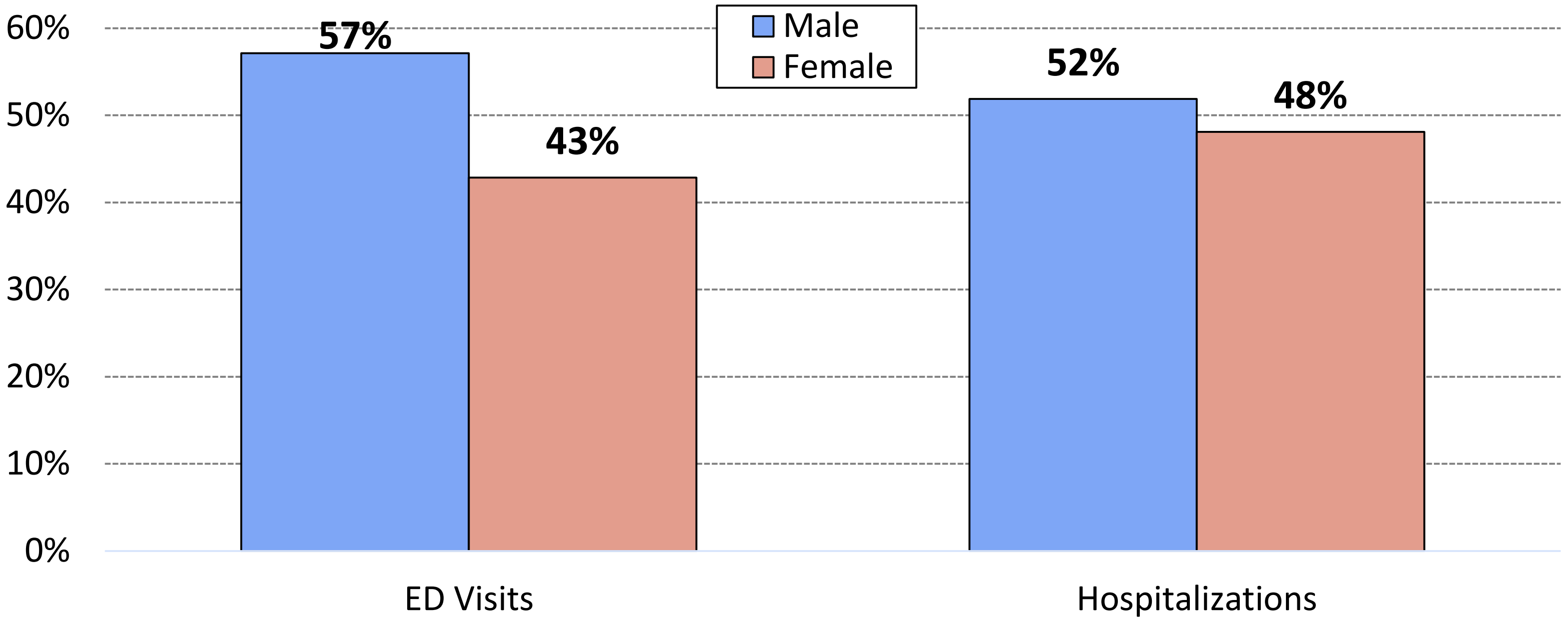
Patient Age Distribution, Opioid-related ED Visits and Hospitalizations



NOTES: Patients with opioid-related ED visits (n= 15,568); Patients with opioid-related hospitalizations (n=24,195). Results are preliminary and the data are not nationally representative.

SOURCE: 2014 NCHS linked to 2014/2015 NDI

Patient Sex Distribution, Opioid-related ED Visits and Hospitalizations



NOTES: Patients with opioid-related ED visits (n= 15,568); Patients with opioid-related hospitalizations (n=24,195)
Results are preliminary and the data are not nationally representative.
SOURCE: 2014 NCHS linked to 2014/2015 NDI

ICD-9-CM Chapter for Primary Diagnosis, Opioid-related ED Visits and Hospitalizations

ICD-9-CM Chapter	ED Visits (n=15,568)	Hospitalizations (n=24,195)
Mental Health Disorders	**44.7%**	**33.8%**
Injury & Poisoning	**20.1%**	**16.7%**
Ill-Defined Conditions	**13.0%**	**3.0%**
Skin Diseases	**3.3%**	3.8%
Musculoskeletal System Diseases	**3.0%**	3.7%
Digestive System Diseases	2.4%	**6.1%**
Complications of Pregnancy	1.8%	**5.8%**
Respiratory System Diseases	2.1%	**5.8%**
Infectious & Parasitic Diseases	0.7%	**5.7%**

Values in **RED** indicate 5 top ranked chapters

Notes: Results are preliminary and the data are not nationally representative.

SOURCE: 2014 NCHS linked to 2014/2015 NDI

Hospital Characteristics for All ED Visits vs. Opioid-related ED Visits

Characteristic	All Visits (n=2,950,103)	Opioid-related Visits (n=15,568)
Northeast	**35.9%**	**46.5%**
Midwest	19.2%	20.8%
South	29.3%	23.7%
West	15.5%	9.1%
Large Central Metro	**48.3%**	**56.0%**
Fringe/Medium/Small Metro	48.4%	42.1%
Rural	3.3%	1.9%
<250 beds	22.6%	22.9%
250-499 beds	**35.3%**	**21.2%**
500-749 beds	**32.3%**	**47.9%**
>=750 beds	9.8%	8.0%

Notes: Results are preliminary and the data are not nationally representative.
SOURCE: 2014 NCHS linked to 2014/2015 NDI

Hospital Characteristics for All Hospitalizations vs. Opioid-related Hospitalizations

Characteristic	All Hospitalizations (n=1,136,045)	Opioid-related Hospitalizations (n=24,195)
Northeast	**35.0%**	**45.8%**
Midwest	15.9%	16.7%
South	32.3%	28.3%
West	16.9%	9.2%
Large Central Metro	45.6%	49.4%
Fringe/Medium/Small Metro	50.9%	47.2%
Rural	3.5%	3.5%
<250 beds	20.2%	23.6%
250-499 beds	**30.2%**	**22.0%**
500-749 beds	**36.4%**	**41.7%**
>=750 beds	13.1%	12.8%

Notes: Results are preliminary and the data are not nationally representative.
SOURCE: 2014 NCHS linked to 2014/2015 NDI

Future Directions

- Completion of the white paper
  - Summary of Task 1 findings analyzing linked data
- Explore other research topics of interest
  - Analyses of other causes of death (e.g., HIV, hepatitis)
  - Comparison of other hospital services received during the hospital visit
  - Cohort study on patients with a history of repeated opioid-related encounters that resulted in an opioid-involved death
  - Other?

Task 2 Update

Enhanced opioid-identification

Task 2: Enhancement of methodologies for identifying opioids

- **Goals**:
  - Enhance existing methods to identify opioid-involvement in linked hospital and mortality data file
  - Develop a new approach based on natural language processing (NLP)
  - Build an enhanced set of algorithms based on all methods to utilize all available structured and unstructured data
- **Timeline**: Present – February 2020

New Staff for Task 2

- Health Research Analyst: Donielle White (11/13)
- NLP Subject Matter Expert: Nikki Adams (11/26)
- FDA ORISE Fellow: Adewumi Adegboye (Early-mid January)

Task 2: Enhancement of methodologies for identifying opioids

- Divided into 5 subtasks:
  - Subtask 2.1: Create an updated 2016 DMI vocabulary.
  - Subtask 2.2: Apply code-based algorithms to identify opioids in 2016 NHCS encounter diagnosis and procedure codes.
  - Subtask 2.3: Apply the DMI program and update vocabulary to 2015 and 2016 NHCS clinical notes and EHR literal text fields.
  - Subtask 2.4: Develop and apply Natural Language Processing (NLP) methodology to the 2016 NHCS clinical notes and EHR literal text fields.
  - Subtask 2.5: Explore the development of a new algorithm that utilizes multiple data sources and methods to identify visits involving specific opioid agents.

Subtask 2.1: Create an updated 2016 Drugs Mentioned with Involvement (DMI) vocabulary

- **Description**: Update the DMI vocabulary using the latest mortality data (2016) to account for new drugs, changes in drug nomenclature, and assign a greater number principal variants.

- **Status**: Complete

Examples of New 2016 DMI Principal Variants

1. AB-CHIMINACA
2. AB-FUBINACA
3. CANNABIDIOL
4. 25C-NBOME
5. 4-FLUROAMPHETAMINE
6. KRATOM
7. ISOBUTYRFENTANYL
8. CARFENTANIL

Subtasks 2.2: Apply code-based algorithms to identify opioids in 2016 NHCS data

- **Timeline**: December 2018 – March 2019
- **Description**: Develop enhanced medical code-based algorithm using following code systems:
  - Diagnosis codes (ICD-10-CM and ICD-10)
  - Procedure codes (CPT, HCPCS, ICD-10-PCS)
  - Medications (RxNORM)
  - Laboratory tests (LOINC)
- **Status**:
  - Currently applying diagnosis and procedure codes, will next identify relevant medication and lab codes
- **Challenges**:
  - Some hospitals only provided SNOMED diagnosis codes, must be converted to ICD-10-CM
  - Algorithm designed for morbidity data, may need to modify to run on mortality data

Subtask 2.3: Apply the updated DMI program to NHCS clinical notes and EHR text fields

- **Timeline**: Present – June 2019
- **Description**: Apply 2016 DMI program with updated vocabulary (subtask 2.1) to hospital clinical notes and other text data to identify opioid-involvement
- **Status**: Currently reviewing associated 2016 DMI SAS programs
- **Challenges**: SAS programs need to be modified to accommodate hospital clinical notes instead of cause of death text.

Subtask 2.4: Develop and apply NLP methodology to NHCS clinical notes and EHR literal text fields

- **Timeline**: May 2018 – September 2019
- **Description**: Identify opioid-involvement in hospital and mortality data text fields using NLP methods
- **Status**:
  - New staff members reviewing available text data and SME will train on NLP basics
  - Create a NLP program that will identify opioid-involvement in hospital EHR data.
- **Challenges**:
  - The clinical notes data are not standard and the type and depth of information varies by hospital.
  - Getting access to necessary software in protected computing environment

Benefits of an NLP-based Approach

- The two major NLP methods, rule-based systems and machine learning, provide a complimentary and efficient way to query both direct and indirect mentions of desired concepts.

- These methods also make it easier to:
  - Capture relevant data despite misspellings, abbreviations, colloquialisms, etc.
  - Add flexibility to queries by adding/modifying new rules as needed.
  - “Teach” the computer to discover new terms and patterns in the data.
  - Examine the context surrounding key terms to help weed out false positives.

Rule-based and Machine Learning NLP Methods

- **Rule-based systems**: human programs a set of “rules” based on specific terms or patterns of terms
  - Example query for drug ingestion—tell machine to look for a specific term: “snort”
- **Machine learning**: human provides machine with basic parameters and general algorithms and then machine uses statistics to “learn” how to categorize text
  - Example query for drug ingestion—build a machine model that understands what words mean, how they are used in context, and have the machine find additional words similar to “snort”: “inhale”, “insufflation”

Opioid NLP Methods: Rule-based systems example

- Compile list of known opioid terms to query clinical notes and cause of death text
  - 2016 DMI vocabulary
  - ESOOS[1] syndromic surveillance search terms

- Query and analyze data to identify and correct misspellings or other variations of opioid terms
  - Example 1—Misspellings: heroin, herion, heroine
  - Example 2—Whitespace variations: Mono Acetyl Morphine vs. Monoacetylmorphine

1. ESOOS = CDC’s Enhanced State Opioid Overdose Surveillance program

Opioid NLP Methods: Machine learning example

Basic Approach to Machine Learning:

- Samples of clinical notes and cause of death literals serve as training sets (corpora) for the machine learning models
- Corpora are human annotated to label various aspects of data, such as:
  - Opioid use implicated in reason for hospital visit (yes/no)
  - Opioid use implicated in cause of death (yes/no)
  - Opioid agent name/category
- Machine will analyze annotated corpora using statistical methods to build its own knowledge, produce its own rules and classifiers

Opioid NLP Methods: Machine Learning example cont.

Using Machine Learning for Opioid-Identification

- Find unknown opioid terms (not just due to spelling variation)

- Program machine to suggest new terms similar to known ones
  - Example—program for fentanyl analogs prompts machine to indicate that “butyr” is like the word “carfentanil”

- This method may be combined with others to quickly detect possible synonyms and automatically categorize them as alternate terms

Cause of Death Example: Applying DMI

Part I - Chain of events that led to death: ACUTE COMBINED DRUG TOXICITY (COCAINE, FURANYL FENTANYL, CARFENTANIL AND HEROIN)

Part II - Other Significant conditions contributing to death: ALPRAZOLAM USE

Box 43 – Describe how the injury occurred: THE DECEDENT USED COCAINE, FURANYL FENTANYL, CARFENTANIL, HEROIN AND ALPRAZOLAM

Cause of Death Example: Applying NLP

Part I - Chain of events that led to death: ACUTE COMBINED DRUG TOXICITY (COCAINE, FURANYL FENTANYL, CARFENTANIL AND HEROIN)

Part II - Other Significant conditions contributing to death: ALPRAZOLAM USE

Box 43 – Describe how the injury occurred: THE DECEDENT USED COCAINE, FURANYL FENTANYL, CARFENTANIL, HEROIN AND ALPRAZOLAM

Hospital Clinical Notes: Applying DMI Program

"Patient PCP: Date: CHIEF COMPLAINT: Unresponsive (Patietn Fount in XXXX bath room on the floor not responding.) HPI: is a 47 year old male with history of diabetes and chronic back pain/chronic opioid use who presents to the ED via EMS with a chief complaint of substance abuse/loss of consciousness. Per EMS was found down in a XXXX bathroom after ""snorting something"". He was unconscious with diminished respiratory rate and pinpoint pupils. He was given intranasal Narcan x 2 and IV Narcan x 1 with arousal. He is currently awake, alert and appropriately responding to questions. Reportedly was in a minor motor vehicle accident yesterday as well as today. After accident yesterday states he developed chest pain which has been persistent throughout the day. He denies head trauma, LOC, or additional injuries. This afternoon he was in a ""3 car fender bender"".

Hospital Clinical Notes: Applying NLP

"Patient PCP: Date: CHIEF COMPLAINT: Unresponsive (Patietn Fount in XXXX bath room on the floor not responding.)   HPI: is a 47 year old male with history of diabetes and chronic back pain/chronic opioid use who presents to the ED via EMS with a chief complaint of substance abuse/loss of consciousness. Per EMS was found down in a XXXX bathroom after ""snorting something"". He was unconscious with diminished respiratory rate and pinpoint pupils. He was given intranasal Narcan x 2 and IV Narcan x 1 with arousal. He is currently awake, alert and appropriately responding to questions.   Reportedly was in a minor motor vehicle accident yesterday as well as today. After accident yesterday states he developed chest pain which has been persistent throughout the day. He denies head trauma, LOC, or additional injuries. This afternoon he was in a ""3 car fender bender"".

Subtask 2.5: Explore the development of a new algorithm that utilizes multiple data sources and methods to identify specific opioid agents

- **Timeline:** April 2019 – February 2020
- **Description:** Harmonize the previous three methods (coded algorithm, DMI program, and NLP) into an enhanced methodology for identifying opioid-involvement
- **Status:** On hold to until completion of Subtasks 2.1-2.4

Task Status Summary

- Task 1: 2014 NHCS/NDI/NVSS-M-DO file
  - Status: On-time
  - Merged 2014 NHCS/NDI/NVSS-M-DO file will be available in the RDC in the early 2019.
  - Report on findings from 2014 NHCS/NDI/NVSS-M-DO will be available in early 2019.
- Task 2: Enhanced opioid-identification in hospital and death certificate data
  - Status: Delayed start due to hiring of NLP expert/analyst.
  - NLP expert and analyst have begun work on opioid-identification.
- Task 3: 2016 NHCS/NDI/NVSS-M-DO file
  - Status: On-time
- Task 4: Dissemination
  - Status: On-time

Discussion

Discussion Questions

- Are there other lists of opioid terms that Nikki should consider for the NLP approach? Perhaps the DEA's list of drug terms?
- Are there NLP researchers at other Federal agencies that we should connect with?