Text Analysis of Death Certificate Records to Ascertain Drugs Involved in Deaths in the National Vital Statistics System

Merianne Rose Spencer[1], Brigham A. Bastian[2], and Margaret Warner[2]

Division of Analysis and Epidemiology[1], Division of Vital Statistics[2]

National Center for Health Statistics

Government Advances in Statistical Programming (GASP) Workshop

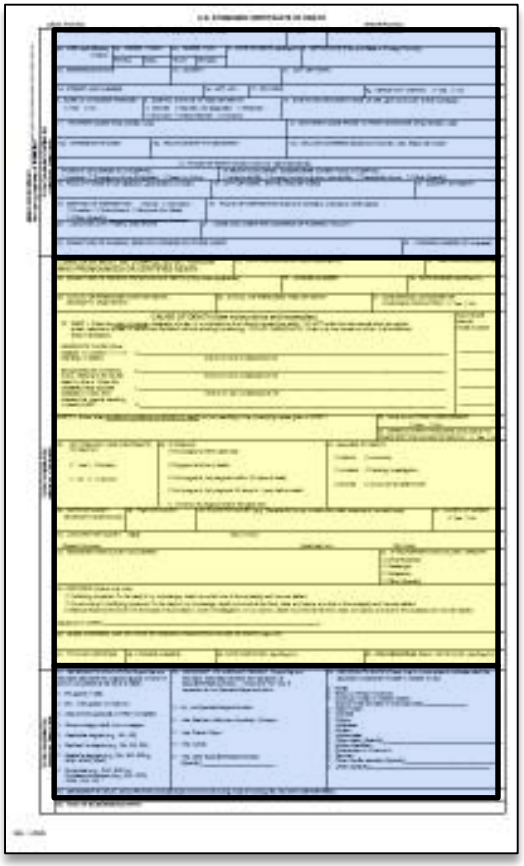
Monday, September 23, 2019

Introduction: NCHS and NVSS

- National Center for Health Statistics (NCHS) provides statistical information that guides actions and policies to improve the health of the American people
- National Vital Statistics System (NVSS) encompasses the processing and coding of birth and death certificate records from 57 jurisdictions*
  - Cause-of-death (ICD-10) codes are assigned to death data received by jurisdictions at NCHS
  - Data are coded, processed and then disseminated for health statistics, surveillance and research

*50 states, New York City, District of Columbia and 5 US territories (American Samoa, Guam, Norther Marianas, Puerto Rico, and Virgin Islands)

Data Source: Death Certificate



Demographic information

Completed by the funeral director using information from *the best qualified person*: spouse, parent, child, another relative, or other person who has knowledge of the facts

Medical information

For natural causes, completed by attending physician, nurse practitioner, physician's assistant

For sudden and unexplained deaths, completed by medical examiner, coroner, Justice of the Peace

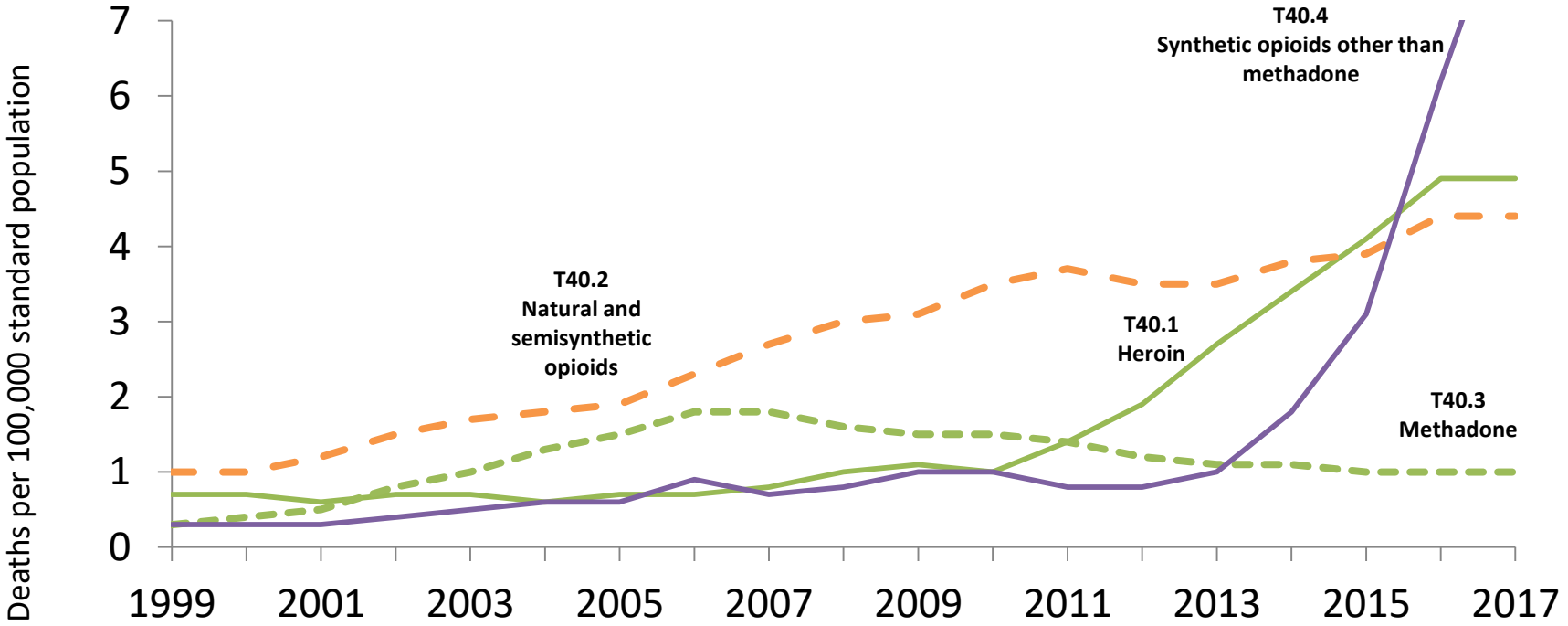
Demographic information

U.S. Standard Death Certificate: Cause-of-Death Section

Referred to as the literal text:

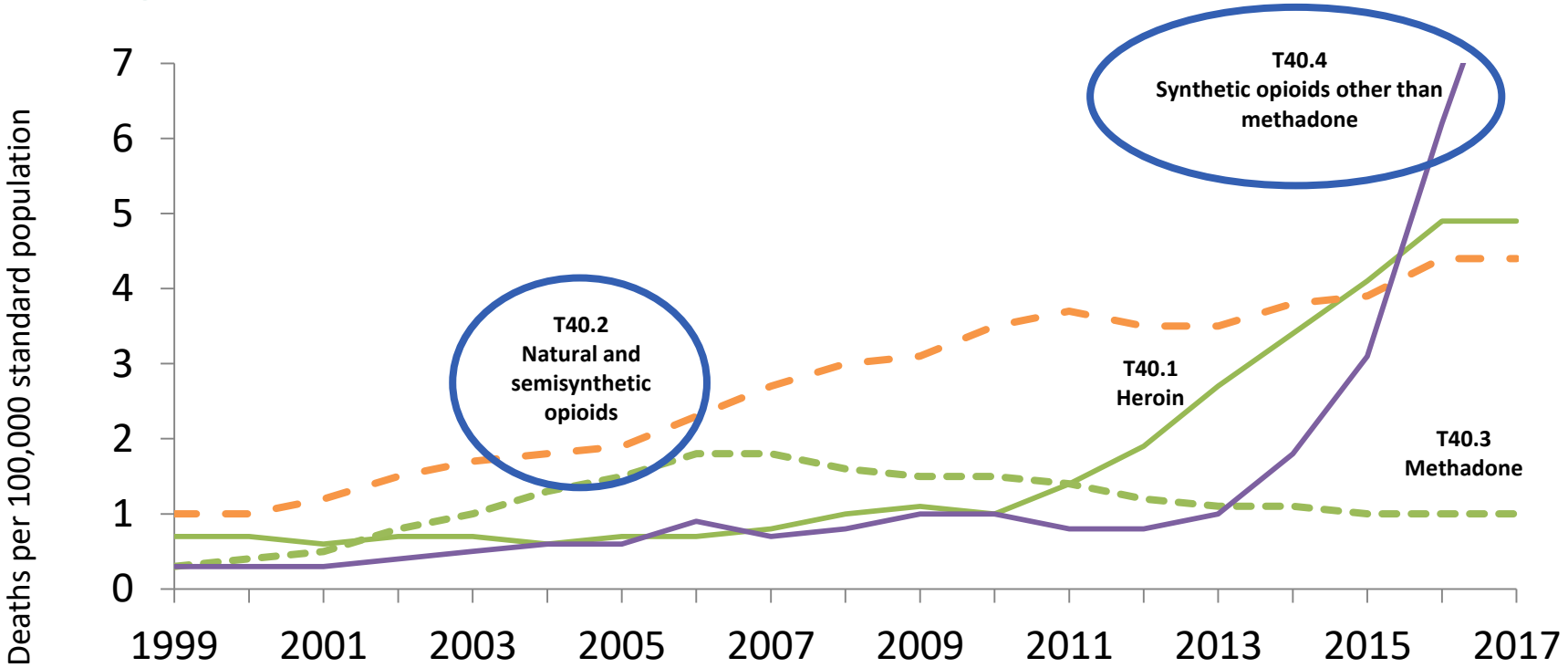
- the information written by the medical certifier on the cause, manner, circumstances, and other factors contributing to the death
  - Part I – Chain of events that directly caused the death
  - Part II – Significant conditions contributing to death
  - Box 43 – Describe how the injury occurred

Drug Overdose Death Rates, United States, 1999-2017



SOURCE: National Center for Health Statistics. National Vital Statistics System Mortality File, 1999-2017.

Drug Overdose Death Rates, United States, 1999-2017



SOURCE: National Center for Health Statistics. National Vital Statistics System Mortality File, 1999-2017.

Examples of common drugs found on death certificates without specific ICD-10 code

Opioids

- Fentanyl
- Fentanyl analogs
- Oxycodone
- Hydrocodone
- Morphine
- Hydrocodone
- Tramadol

Stimulants

- Methamphetamine
- MDMA

Benzodiazepines

- Alprazolam
- Diazepam
- Clonazepam

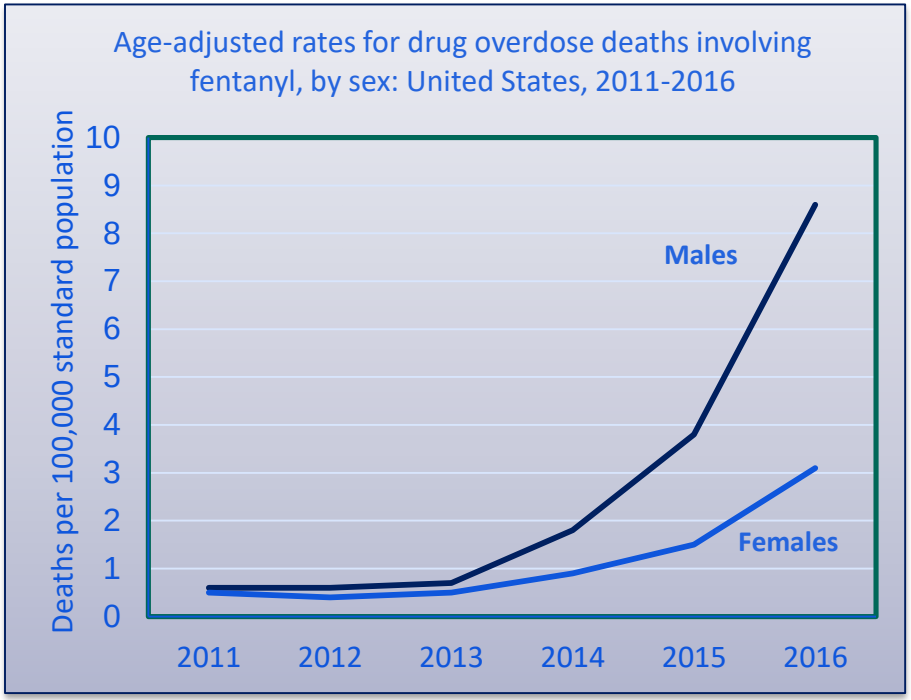
3-Methylfentanyl	Despropionyl-fentanyl
4-ANPP	Furanyl fentanyl
4-MeO-Butyrylfentanyl	Methylfentanyl
Acetylfentanyl	Para-fluorobutyrl-fentanyl
Acrylfentanyl	Para-fluoroisobutyrylfentanyl
Acryloylfentanyl	P-Fluoro-butyrylfentanyl
Alfentanil	Remifentanil
Butyrylfentanyl	Sufentanil
Carfentanil	…

Methods to Identify Specific Drugs on Death Records

- Collaboration between NCHS and FDA
- Developed methods to analyze literal text for mentions of specific drugs
  - Preprocessed text (e.g., remove stop words, special characters)
  - Examined and reviewed string terms and phrases
  - Categorized terms into categories (i.e., principal variants)
  - Considered context (e.g., “history of”, “insulin-dependent”)
- Referred to as the Drugs Mentioned with Involvement (DMI) methodology[1]

[1] Drugs Mentioned with Involvement (DMI) methodology : https://www.cdc.gov/nchs/data/nvsr/nvsr65/nvsr65_09.pdf

Selected Results



SOURCE: Spencer MR, Warner M, Bastian BA, et al. Drug overdose deaths involving fentanyl, 2011-2016. Natl Vital Stat Rep. 2019 Mar;68(3):1-19.

Top drugs involved in drug overdose deaths, 2011 and 2016

	2011		2016	
Rank	Referent drug	# of deaths	Referent drug	# of deaths
1	Oxycodone	5,587	Fentanyl	18,335
2	Cocaine	5,070	Heroin	15,961
3	Heroin	4,571	Cocaine	11,316
4	Methadone	4,545	Methamphetamine	6,762
5	Alprazolam	4,066	Alprazolam	6,209
6	Morphine	3,290	Oxycodone	6,199
7	Hydrocodone	3,206	Morphine	5,014
8	Methamphetamine	1,887	Methadone	3,493
9	Diazepam	1,698	Hydrocodone	3,199
10	Fentanyl	1,662	Diazepam	2,022

SOURCE: Hedegaard H, Bastian BA, Trinidad JP, Spencer MR, Warner M. Drugs most frequently involved in deaths: United States, 2011-2016. Natl Vital Stat Rep. 2018 Dec;67(9):1-14.

Considerations and Next Steps

- Literal text analyses can enable researchers to report with more granularity, the specific drugs involved in deaths
- Challenges
  - Identification of specific drugs is dependent on testing, interpretation and reporting by death certifiers
  - Temporal changes in reporting of the specific drugs can impact trends analyses
- Future steps
  - Automate these processes (e.g., machine learning), while recognizing that we must also identify terms that appear a small number of times

Questions?

Merianne Rose T. Spencer, M.P.H.
Epidemiologist
Division of Analysis and Epidemiology
National Center for Health Statistics
301-458-4377 | kvd1@cdc.gov

Brigham A. Bastian, B.S.
Statistician (Demography)
Division of Vital Statistics
National Center for Health Statistics
301-458-4517 | xcb7@cdc.gov

Margaret Warner, Ph.D.
Injury Epidemiologist
Division of Vital Statistics
National Center for Health Statistics
301-458-4556 | mmw9@cdc.gov