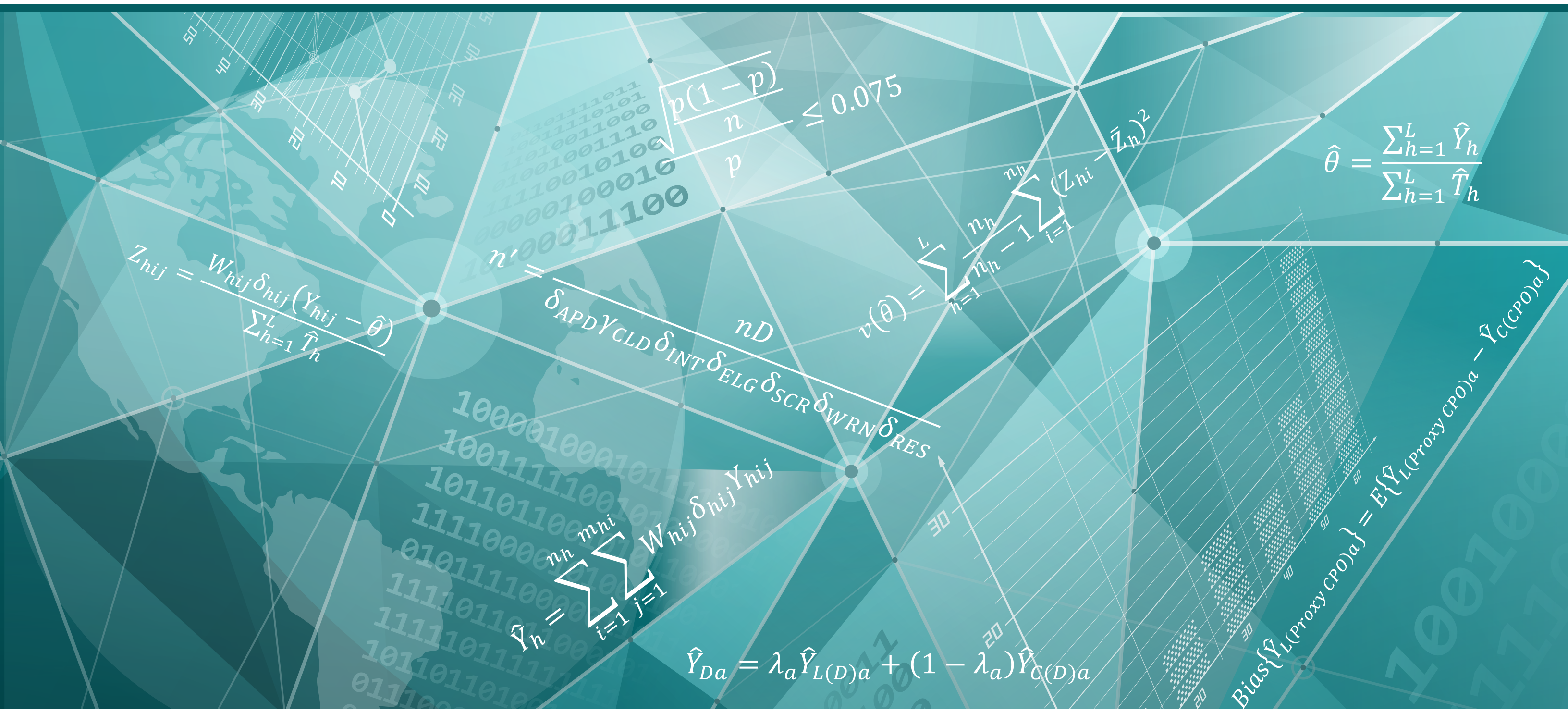
NATIONAL CENTER FOR HEALTH STATISTICS

Vital and Health Statistics

Series 2, Number 188 October 2021



Enhancing Identification of Opioid-involved Health Outcomes Using National Hospital Care Survey Data

Data Evaluation and Methods Research

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

NCHS reports can be downloaded from: https://www.cdc.gov/nchs/products/index.htm.

Suggested citation

White DG, Adams NB, Brown AM, O'Jiaku-Okorie A, Badwe R, Shaikh S, Adegboye A. Enhancing identification of opioid-involved health outcomes using National Hospital Care Survey data. National Center for Health Statistics. Vital Health Stat 2(188). 2021.
DOI: https://dx.doi.org/10.15620/cdc:108998.

For sale by the U.S. Government Publishing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20401–0001
Printed on acid-free paper.

NATIONAL CENTER FOR HEALTH STATISTICS
Vital and Health Statistics

Series 2, Number 188 October 2021

Enhancing Identification of Opioid-involved Health Outcomes Using National Hospital Care Survey Data

Data Evaluation and Methods Research

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
October 2021

Contents

Text Figures

Table

Detailed Tables

Contents—Con.

Appendix Tables

Enhancing Identification of Opioid-involved Health Outcomes Using National Hospital Care Survey Data

by Donielle G. White, M.P.H., Nikki B. Adams, Ph.D., Amy M. Brown, M.P.H., Adaeze O'Jiaku-Okorie, M.P.H., Rihem Badwe, Pharm.D., Salah Shaikh, M.P.H., and Adewumi Adegboye, M.P.H.

Abstract

Purpose

This report documents the development of the 2016 National Hospital Care Survey (NHCS) Enhanced Opioid Identification Algorithm, an algorithm that can be used to identify opioid-involved and opioid overdose hospital encounters. Additionally, the algorithm can be used to identify opioids and opioid antagonists that can be used to reverse opioid overdose (naloxone) and to treat opioid use disorder (naltrexone).

Methods

The Enhanced Opioid Identification Algorithm improves the methodology for identifying opioids in hospital records using natural language processing (NLP), including machine learning techniques, and medical codes captured in the 2016 NHCS. Before the development of the Enhanced Opioid Identification Algorithm, opioid-involved hospital encounters were identified solely by coded diagnosis fields. Diagnosis codes provide limited information about context in the hospital encounters and can miss opioid-involved encounters that are embedded in free text data, like hospital clinical notes.

Results

In the 2016 NHCS data, the enhanced algorithm identified 1,370,827 encounters involving the use of opioids and selected opioid antagonists. Approximately 20% of those encounters were identified exclusively by the NLP algorithm.

Keywords: opioids • health care • hospitals

Introduction

In fiscal year 2018 (FY18), the Office of the Secretary-Patient Centered Outcomes Research Trust Fund (PCORTF) awarded the National Center for Health Statistics (NCHS) funding to develop the Enhanced Opioid Identification Algorithm to improve the identification of opioid-involved hospital encounters in the National Hospital Care Survey (NHCS). NHCS assesses the health of the population by collecting information on health care utilization as well as the demographic characteristics, medical conditions, and treatment of patients who use hospitals for inpatient and ambulatory medical care in the United States, although currently the data are not nationally representative. The enhanced algorithm identifies opioid-involved hospital encounters using all available data fields in hospital billing and electronic health record (EHR) data collected in NHCS. Comprehensive data on opioid-involved emergency department (ED) encounters and inpatient hospitalizations will assist researchers in identifying and testing strategies to reduce morbidity and mortality associated with the misuse and overdose of opioids.

Before this project, opioid-involved hospital encounters in NHCS could only be captured from information in the coded diagnosis fields. The Enhanced Opioid Identification Algorithm seeks to improve the methodology for identifying opioids in hospital records by using natural language processing (NLP) techniques, including machine learning, in addition to the medical codes captured in the 2016 NHCS. NLP techniques use linguistics and computer science methods to process and categorize a large amount of human language data into a format that can be easily analyzed. The medical codes captured in the 2016 NHCS include diagnosis, procedure, medication, and laboratory test result codes. The NLP portion of the algorithm is designed to extract information from the clinical notes associated with the EHR portion of NHCS. The output from the Enhanced Opioid Identification Algorithm, the Enhanced Opioid Identification Data Set, is available to researchers through the Federal Research Data Centers (RDC). Please see the RDC website (https://www.cdc.gov/rdc/index.htm) for instructions on submitting a proposal.

This report details the methodology used to create the Enhanced Opioid Identification Algorithm, starting with the background of the project. Next, the case definitions of opioid involvement, opioid overdose, and specific opioids mentioned are discussed. The methodology used in the development of the Enhanced Opioid Identification Algorithm is then described, followed by the results of the algorithm. Analytic considerations and limitations of the enhanced methodology are detailed, followed by future considerations and uses of the algorithm.

Project Background

The FY18 PCORTF project was developed to improve the identification of opioids mentioned before or during a hospital encounter in the 2014 and 2016 NHCS data. Hospitals participating in the 2014 NHCS could only submit Uniform Billing (UB)–04 administrative claims data. The UB–04 data classified diagnosis information exclusively using *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD–9–CM) diagnosis codes (1). ICD–9–CM codes provide limited information about specific drug agents (opium, heroin, and methadone) but do capture information on novel drugs (illicit fentanyl analogs). The inability to identify specific opioids highlighted the limitations of relying exclusively on diagnosis codes and the need for new strategies to identify opioid-involved hospital encounters (2).

In the 2016 NHCS, in addition to submitting UB–04 administrative claims data, participating hospitals had the option to submit EHR data or data collected by Vizient, a large provider-driven, health care performance improvement organization. The addition of Vizient and EHR sources presented the opportunity to identify opioid-involved hospital encounters using data elements not available in UB–04 administrative claims data (medication data, laboratory data, and clinical notes).

The 2014 and 2016 NHCS provided an opportunity to address two main aims of the FY18 PCORTF project: a) to identify opioid-involved hospital encounters using ICD–9–CM codes and b) to expand the search of opioid identification beyond diagnosis codes. To meet the second aim, the Enhanced Opioid Identification Algorithm was developed to identify opioid-involved hospital encounters, opioid overdose hospital encounters, and 17 specific opioid drug categories in the 2016 NHCS hospital data.

The UB–04 administrative claims data contain information on patient demographics, identifiers, conditions, services, and discharge status. EHR data include similar data available in UB–04 administrative claims as well as additional items that provide more detail about a patient's hospital encounter, including medications, clinical notes, and laboratory results. Diagnosis codes are often missing from EHR records but information on diagnoses is often available in EHR text data fields. Vizient collects UB–04 administrative claims and obtains data on medications and laboratory tests. Missing diagnosis codes were present in text fields in the EHR data, allowing researchers to enhance the identification of opioid-involved hospital encounters with NLP and machine learning data analysis techniques.

Data Source

NHCS is designed to produce national estimates on the characteristics of inpatient hospitalizations and ED encounters, including length of stay of inpatient encounters, diagnoses, surgical and nonsurgical procedures, and patterns of hospital utilization in various regions of the country (3). The target universe for NHCS is all inpatient discharges and in-person visits made to EDs in noninstitutional, nonfederal hospitals in the 50 states and the District of Columbia that have six or more staffed inpatient beds. Data are extracted from hospital billing or EHR systems and then transmitted electronically directly to NCHS or its designated agent. The 2016 NHCS collected EHR data in two formats: custom extracts and continuity of care documents (CCD). In 2016, the NHCS sample included 581 hospitals. The survey also collects patient personally identifiable information (PII), which allows researchers to both follow patients who have multiple hospital encounters and link patients to external data sources such as the National Death Index (NDI). However, Vizient does not include PII on the file, so linkage to external data sources is not possible.

In the 2016 NHCS, 158 hospitals submitted data: 89 hospitals submitted UB–04 administrative claims data, 16 hospitals submitted custom extracts of EHR data, 31 hospitals submitted EHR data in the form of CCD, and 22 hospitals submitted data via Vizient. The Enhanced Opioid Identification Algorithm was applied to 9,624,026 ED and inpatient encounters in the 2016 NHCS. However, while intended to make national estimates, neither the 2014 nor the 2016 NHCS provide nationally representative data due to low response rates (16% for 2014 and 27% for 2016).

Case Definitions

Case definitions were developed to identify opioid involvement, opioid overdose, and specific opioid(s) mentioned. The case definitions were developed in collaboration with a Technical Expert Panel (TEP) and the NCHS Board of Scientific Counselors (BSC) PCORTF Drug Work Group (https://www.cdc.gov/nchs/data/bsc/bsc_mintues_september_2019.pdf). The TEP included representatives from several federal agencies, including NCHS, Food and Drug Administration (FDA), Substance Abuse and Mental Health Services Administration (SAMHSA), and National Institute on Drug Abuse. The NCHS BSC PCORTF Drug Work Group included representatives from universities and public health institutions. The TEP and NCHS BSC PCORTF Drug Work Group members provided subject matter expertise on identifying and classifying opioids, as well as medical codes and search terms that could be used to identify opioid-involved encounters.

Opioid-involved Encounters

An opioid-involved encounter was defined as one mentioning past or present use of an opioid. The term “involved” includes mentions of any form of opioid use before arrival at the hospital (previously prescribed opioids or previously taken illicit opioids) and mentions of opioids administered during the encounter or prescribed upon discharge. A hospital encounter identified as opioid-involved could have a mention of opioid use in the EHR clinical notes or a medical code associated with an opioid.

Opioids included the following substances:

- Natural opioids (morphine and codeine)
- Semisynthetic opioids (oxycodone and hydrocodone)
- Prescription synthetic opioids (tramadol and fentanyl)
- Illicit opioids (heroin, krokodil, and illicitly manufactured fentanyl and its analogs)

The following forms of opioid use were included:

- *Prescribed use*—Taking an opioid as prescribed or directed.
- *Misuse*—Use of illegal opioids or the use of prescription opioids in a manner other than as directed by a doctor, such as use in greater amounts, more often, or for longer than told to take a drug or using someone else’s prescription.
- *Opioid use disorder*—A problematic pattern of opioid use that causes significant impairment or distress. A diagnosis is based on specific criteria such as unsuccessful efforts to cut down or control use, or use resulting in social problems and failure to fulfill obligations at work, school, or home, among other criteria.
- *Overdose*—Taking an opioid in an excessive amount, either intentionally or unintentionally, that causes injury to the body (poisoning).
- *Adverse effects*—When an opioid intended for therapeutic use has an unintended and injurious effect.
- *Underdosing*—Taking less of a prescription opioid than is prescribed by a provider or a manufacturer’s instruction.
- *Miscellaneous*—Other forms of opioid use that can be identified in the *International Classification of Diseases, 10th Revision, Clinical Modification* (ICD–10–CM) coding system, including chronic opioid analgesic use, newborns affected by maternal use of opioids, and presence of opioids in blood (4–6).

Three criteria were established to meet the opioid-involved case definition:

1. Presence of at least one selected opioid use code in any diagnosis, reason for visit, problem, procedure, or medication code field;
2. Evidence of a positive laboratory test indicating presence of an opioid; or
3. Classification by the NLP processor based on opioid use indicators in the text clinical notes.

Opioid Overdose Encounters

The opioid overdose case definition was created to identify a subset of opioid-involved encounters mentioning opioid toxicity or poisoning due to ingesting a high dose of opioids. Two criteria were established for opioid overdose encounters:

1. Presence of at least one selected opioid overdose code indicating poisoning or acute intoxication in any diagnosis, reason for visit, or problem code field; or
2. Classification by the NLP processor based on opioid overdose indicators in the text clinical notes.

In previous studies, naloxone administration has been examined as a proxy indicator for an overdose event (7,8). However, recent guidance has recommended that naloxone be administered for all suspected overdoses, even when the substance taken is unknown, and therefore is not always administered exclusively for the ingestion of opioids (9). Also, in the hospital setting, a naloxone-containing product such as prolonged-release oxycodone or naloxone may be administered to manage postoperative pain rather than to reverse opioid toxicity (10–12). Therefore, a naloxone mention was insufficient to meet the opioid overdose case definition.

Specific Opioids Mentioned

All mentions of the specific types of opioids taken by the patient were identified and assigned to 17 drug categories listed in the Table.

The first 15 categories represent 13 commonly used opioids and 2 opioid antagonists that can be used to reverse opioid overdose (naloxone) and to treat opioid use disorder

Table. Drug categories identified by the Enhanced Opioid Identification Algorithm

Drug category
1. Buprenorphine or norbuprenorphine
2. Codeine
3. Fentanyl or fentanyl analogs
4. Heroin (6–AM and 6–MAM)
5. Hydrocodone
6. Hydromorphone
7. Levorphanol
8. Meperidine
9. Methadone
10. Morphine
11. Naloxone
12. Naltrexone
13. Oxycodone
14. Oxymorphone
15. Tramadol
16. Other opioid
17. Unspecified opioid

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

(naltrexone). The inclusion of opioid antagonists will enable researchers to explore the use of these pharmacological treatments among encounters flagged for opioid involvement and opioid overdose. Mentions of specified opioids other than the 15 named opioids or opioid antagonists were categorized as "other opioid" and each specific drug name was captured in a separate comma delimited field in the Enhanced Opioid Identification Data Set. Mentions of unspecified opioids were categorized as "unspecified opioid." Lastly, mentions of drug combinations containing multiple opioids or opioid antagonists were assigned to all categories that applied. For example, mentions of the medication "buprenorphine or naloxone" were assigned to categories for "buprenorphine or norbuprenorphine" and "naloxone."

Three criteria were developed to identify specific opioids used:

1. Presence of at least one selected code specifying type of opioid used in any diagnosis, reason for visit, problem, procedure, or medication code field;
2. Evidence of a positive laboratory test indicating presence of a specific opioid or unspecified opioids; or
3. Classification by the NLP processor based on indicators for type of opioid used in the text clinical notes.

Case Definition Medical Codes and Search Terms

The methodology used to identify medical codes and search terms and to build NLP processors for each case definition is described in the Enhanced Opioid Identification Methodology section. Final medical code and search term lists can be found in the NCHS RDC file specifications (https://www.cdc.gov/nchs/data/nhcs/Task-3-Doc-508.pdf, see Appendices I–VI).

Enhanced Opioid Identification Methodology

The Enhanced Opioid Identification Algorithm consists of two components. The first component uses data associated with medical codes and the second component uses NLP techniques on the literal text fields. The two-component approach allowed for an efficient method of identifying opioid-involved encounters and used all available information collected in NHCS.

Code Component Development

In the development of the code component, lists of medical codes and search terms were developed to identify opioid involvement, opioid overdose, and specific opioids taken using a two-phased approach. The lists of medical codes included diagnosis and service codes. In the first phase, relevant codes and terms were extracted from existing lists provided by federal government and academic sources. In the second phase, the study team further refined each list to match the inclusion and exclusion criteria of the final case definitions. This process is described in more detail below.

Code component phase 1: Initial code and search term lists

In the initial phase, the study team collaborated with TEP to identify existing lists of medical codes and search terms from the following entities:

- NCHS' Division of Health Statistics and SAMHSA—Medical codes used in an earlier set of algorithms to identify substance-involved emergency department visits in NHCS 2014 (3).
- NCHS' Division of Vital Statistics (DVS) and FDA—Drugs involved in mortality (DIM) search terms for licit and illicit drugs, drug classes, and drug exposures not otherwise specified to identify drug mentions on death certificates (5).
- Centers for Disease Control and Prevention's (CDC) National Center for Injury Prevention and Control—Suspected opioid overdose terms and medical codes developed for the Enhanced State Opioid Overdose Surveillance (ESOOS) program to monitor overdose trends in participating states (13).
- CDC, FDA, and the Consumer Product Safety Commission—Selected pharmaceuticals to code adverse drug events in U.S. hospital emergency departments for the National Electronic Injury Surveillance System–Cooperative Adverse Drug Event Surveillance (NEISS–CADES) project (6).
- Drug Enforcement Agency—Drug slang terms and code words to assist law enforcement personnel in identifying a wide variety of controlled substances (14).
- University of Kentucky—Drug terms used to search free-text fields in death certificates for the state of Kentucky before ICD–10 coding by NCHS DVS (15).

All relevant information was extracted from each existing list and compiled for review by study team members trained in pharmacology and emergency medicine and additional subject matter experts from the NCHS Clinical Advisory Group.

Code component phase 2: Refining code and search term lists

In the second phase, diagnosis and service codes were analyzed to identify opioid-involved and opioid overdose encounters. The service codes included information on procedures, medications, and laboratory test results. Also, initial diagnosis and service codes and search terms were refined to ensure comprehensiveness and to meet all aspects of the case definitions.

Code component: Diagnostic medical codes

The original opioid-identification algorithm used in Task 1 of this project used selected ICD–9–CM diagnosis codes to identify substance involvement in the 2014 NHCS data. In 2016, participating hospitals submitted diagnostic information in multiple coding systems but primarily using the newer ICD–10–CM. Relevant codes from the original algorithms indicating opioid involvement were therefore mapped to the equivalent ICD–10–CM codes. Additional categories of ICD–10–CM codes were added to the case definition to cover all forms of opioid use. This included codes for opioid use, abuse, dependence, adverse events, poisoning, and underdosing. All diagnostic code fields in 2016 NHCS data were analyzed to identify opioid-involved visits, including diagnosis, reasons for visit, and problems.

Hospitals that submitted EHR data in the 2016 NHCS had diagnostic information submitted using the Systematized Nomenclature of Medicine–Clinical Terms (SNOMED–CT) code system. SNOMED–CT codes were mapped to the equivalent ICD–10–CM codes; those that were not able to map to ICD–10–CM were left in the original code system. Table 1 shows that in the 2016 NHCS data, there were 6,988,635 ED and 2,556,606 inpatient encounters with at least one reported ICD–10–CM code in a diagnosis, reason for visit, or problem list coded field. A relatively small percentage of encounters (0.6% ED and 1.4% inpatient) did not have a diagnosis record with an ICD–10–CM code. The diagnostic code fields of the encounters without an ICD–10–CM code were not searched, but their service code fields (procedures, medications, and laboratory tests) were searched for terms indicating opioid involvement and specific types of opioids used to identify opioid-involved encounters.

Service medical codes

The substance-involvement algorithms used to analyze the 2014 NHCS included service codes in several different coding systems:

- ICD–9–Procedure Classification System (PCS);
- Current Procedural Terminology (CPT); and
- Healthcare Common Procedure Coding System (HCPCS).

To adapt the original service code list to 2016 NHCS data, ICD–9–PCS codes related to opioid use were mapped to equivalent codes in the newer 10th revision coding system, ICD–10–PCS. In addition, the original list of CPT and HCPCS codes were checked for updates released in 2016 and revised to reflect pertinent additions and deletions. Participating hospitals could also submit service-related codes in the following systems:

- SNOMED–CT;
- RxNorm; and
- Logical Observation Identifiers Names and Codes (LOINC).

A list of applicable SNOMED–CT and RxNorm codes was developed and refined using research tools to ensure inclusion of codes for all generic medications (active ingredient) and brand names, illicit opioids, and alternative misspellings. The regulatory status of each medication and whether it was branded in the United States or globally were checked. Resources included the Micromedex website in addition to multiple references such as Clinical Pharmacology, Elsevier, and the American Hospital Formulary Service (16,17). Substances that were not opioids or did not induce opioid-like effects were excluded from the list.

Finally, LOINC codes were used to identify laboratory tests to detect the presence of opioids and metabolites in the body. Resources such as the U.S. National Library of Medicine's Value Set Repository were reviewed to ensure inclusion of screening and confirmatory tests for all opioid-related components across all systems of the body (18).

Service search terms

Some EHR hospitals provided vendor-supplied or hospital-specific service codes that could not be translated to a standard coding system. However, if there was an accompanying descriptive label, these encounters were analyzed for the presence of search terms for medications, procedures, or laboratory tests.

Overall, the availability of service code and label entries that could be searched was greater for procedure fields compared with laboratory test and medication fields. While 89.9% of ED encounters and 92.9% of inpatient encounters had at least one procedure code or label that could be searched, only one-quarter (25.6%) of ED and more than one-third (37.5%) of inpatient encounters had a medication code or label that could be searched (Tables 2–4). Only 7.2% of ED encounters and 15.7% of inpatient encounters had a laboratory code or label that could be searched.

Conducting search for code component

The revised medical code and search term lists from the second phase were then used to search all available diagnostic and service code fields in the 2016 NHCS. SAS 9.4 was used to perform all code-based searches and both SAS 9.4 and Python 9.4 were used to perform all text-based searches of code labels (code descriptions). The final medical code and search term lists can be found in the NCHS RDC file specifications (https://www.cdc.gov/nchs/data/nhcs/Task-3-Doc-508.pdf, see Appendices I–VI).

Identification of opioid-involved encounters

The following searches were conducted to identify each encounter's opioid involvement:

- Diagnosis, reason for visit, and problem code fields for selected ICD–10–CM codes;
- Procedure code fields for selected ICD–10–PCS, HCPCS, and CPT codes;

- Procedure code labels (if nonstandard code was present) for specific opioids;
- Medication code fields for selected SNOMED–CT and RxNorm codes;
- Medication code labels (if nonstandard code was present) for specific opioids;
- Laboratory code fields for selected LOINC codes; and
- Laboratory code labels (if nonstandard code was present) for specific opioids.

Searches for laboratory tests required a two-part process. If a LOINC code or opioid drug was found in a laboratory code or label field, then a keyword search was performed on the accompanying qualitative test result to determine the presence of an opioid in the tested specimen, (a positive result as indicated by keywords such as a "positive" and "pos"). Laboratory test results with a positive result met inclusion criteria as opioid-involved, while laboratory tests with either no results or a negative result were excluded.

Identification of opioid overdose encounters

All diagnostic fields (diagnosis, reason for visit, and problem) were searched for a subset of ICD–10–CM codes indicating opioid poisoning and opioid use, abuse, or dependence with intoxication. Opioid overdose encounters with visit dates before or after calendar year 2016 were not searched by the code component of the enhanced algorithm. This exclusion was intended to limit the search to encounters most likely to involve patients who experienced an acute overdose before presenting at the hospital during calendar year 2016.

Identification of specific opioids mentioned

The following searches were conducted to identify the specific opioid(s) mentioned for each encounter:

- Diagnosis, reason for visit, and problem code fields for a subset of ICD–10–CM codes indicating use of specific opioids;
- Procedure code fields for selected ICD–10–PCS, HCPCS, and CPT codes;
- Procedure code labels (if nonstandard code was present) for specific opioids;
- Medication code fields for selected SNOMED–CT and RxNorm codes;
- Medication code labels (if nonstandard code was present) for specific opioids;
- Laboratory code fields for selected LOINC codes; and
- Laboratory code labels (if nonstandard code was present) for specific opioids or antagonists.

Laboratory tests for an opioid must have been accompanied by evidence of a positive test result. Specific opioid mentions were assigned to all applicable categories for the 15 named opioids or opioid antagonists, other specified opioids, and unspecified opioids as noted in the Case Definitions.

If medication codes were not available, medication labels were searched for opioid-involved medications. Some medication labels had information about the medication name (brand or generic), dosage, route of administration, or a combination of all three. Some hospitals also provided the medication status (if a medication was given during the hospital encounter). Opioid-involved medications with a status that confirmed the medication was administered during the encounter were coded as opioid-involved (given or cosigned). If an opioid-involved medication had a negation term (not given, error, or aborted) or if the medication was given during a clinical trial (study drug or placebo) the medication was not captured by the algorithm.

Combination opioid medications are captured by the code component of the Enhanced Opioid Identification Algorithm. Drug combinations that contained both an opioid antagonist or agonist were mapped to their respective categories. For example, if an encounter mentioned the medication "buprenorphine or naloxone," the categories for "buprenorphine" and "naloxone" are positive for the encounter.

NLP Component Development

Previous algorithms developed for NHCS data, which relied exclusively on medical codes, may miss some cases and lead to an undercount of the true number of substance-involved encounters (3). Text-rich elements in clinical notes may contain additional evidence regarding the use of substances like opioids that would otherwise be missed by purely code-based algorithms. The NLP component used text mining techniques and machine learning to efficiently search and categorize all available clinical notes into a format that allows for classification of opioid-involved or overdose encounters. The use of similar NLP methods has been previously applied to assess problematic opioid usage, inappropriate opioid-related behaviors, and opioid overdoses (19,20).

Annotation of gold standard data set

To test the performance of the NLP component, a gold standard data set of 2016 NHCS encounters identified as opioid-involved and opioid overdoses was needed. This gold standard data set was created by a team of three clinically trained annotators annotating 2,000 inpatient and ED encounters. The annotation involved manually reviewing all available clinical data collected for those hospital encounters in NHCS. The data reviewed included diagnostic and service medical codes or labels and text-rich elements (clinical notes). An annotation guide and form were developed before the annotation to standardize the annotation process and ensure that each annotator was reviewing each case similarly. Input from TEP members was incorporated into the development of the annotation form and guide.

The goal of annotation was to ensure that a sufficient number of encounters was selected to test the performance of NLP processors for each of the three case definitions: opioid involvement, opioid overdose, and specific opioids mentioned.

Before the annotation of 2,000 encounters, a pretest was conducted to refine the annotation form and ensure that each annotator agreed on which encounters were opioid-involved. The pretest encounters included EHR data submitted as custom extracts and CCDs with clinical notes containing opioid-involved ICD–10–CM diagnosis codes or keywords, or neither. There were four rounds of pretesting involving annotating a random sample of encounters. The annotators applied the annotation guide and form to assist them in identifying opioid-involved encounters and to ensure the gold standard annotation data set efficiently and accurately captured the opioid-involved outcomes to build and test the NLP algorithm and record their responses in a database. Throughout the pretest process, the clinically trained annotators assessed the annotation guide and form. During the pretest annotation, inconsistencies and disagreements between annotators were discussed and reconciled. Annotators' responses were evaluated for inter-annotator reliability (also known as inter-rater agreement) as a measure of reliability. Specifically, inter-annotator reliability assesses the extent to which two or more annotators agree on the annotation of data when they perform the annotation independently of one another (21).

The pretest was completed after the inter-annotator agreement among the three clinicians was over 90% on the most critical questions identifying opioid involvement, and then full annotation began. Upon completion, the clinicians met to review, discuss, and adjudicate by consensus certain cases where there was disagreement in the classification of the FY18 and FY19 categories. Additionally, quality control checks were run on annotation responses to identify potential data entry errors from annotators.

The full annotation, consisting of a stratified, random sample of approximately 2,000 2016 NHCS encounters, was drawn across nine categories for the annotation. The eligibility criteria for the categories were not mutually exclusive. However, each encounter was counted as belonging to only one category (after an encounter was selected for a category, the encounter was no longer available for selection in another category even if it met the inclusion criteria). Descriptions of the categories for the FY18 project and the number of encounters selected for each are detailed below:

1. Opioid-relevant medical codes: Encounters whose ICD–10–CM diagnosis codes indicate opioids are involved. Encounters selected: 100.
2. Opioid terms: Encounters where relevant text fields contain an opioid term, determined by a keyword search, excluding encounters that had an opioid-relevant diagnosis code. These terms may come from the Electronic Surveillance System for the Early Notification of Community-based Epidemics query used by the opioid project at ESOOS, terms for drugs in the Drug Mentioned with Involvement (DMI) program used by DVS, or terms from the NEISS–CADES program. From the complete set of all encounters that contain an opioid term, any encounters that also had a relevant ICD–10–CM diagnosis code were omitted. Encounters selected: 500.
3. Opioid overdose codes: Encounters that have an opioid overdose ICD–10–CM diagnosis code. Encounters selected: 50.
4. Opioid overdose keywords: Encounters that have keyword matches for opioid overdose, but do not have an ICD–10–CM overdose diagnosis code. Encounters selected: 200.
5. Additional encounters: Encounters included for the development of the FY19 PCORTF project algorithm identifying substance use disorder (SUD) and mental health issues (MHI). Encounters selected: 800.
6. A random selection: Implemented to balance the data set (ensures that negative examples are available for evaluation) and may include some cases that are relevant, but did not fit into any of the above categories of encounters. The set from which these are selected excludes encounters that had a relevant ICD–10–CM diagnosis code (opioid-involved, SUD-related, or MHI-related) as well as any that had an opioid term, MHI keyword term, or SUD term. Encounters selected: 300.

The final annotation data set was used as a developmental and evaluation data set for the NLP processors, and those outcomes are reported in Validating and Refining NLP Processors Against Gold Standard Annotated Data.

Identifying opioid involvement and specific opioids mentioned

Per the case definition, opioid involvement meant past or present use of an opioid. The goal of the NLP component was to identify and classify opioid-involved encounters by performing the following tasks:

- Perform upfront exclusions (described below);
- Find opioid search terms;
- Detect rule-outs (negated terms like "did not use" and dates beyond the scope of the survey [after 2016]); and
- Assign specific opioid mentions to the 17 drug categories of interest.

Based on early analysis of false positives, upfront exclusions were performed on any note categorized as patient instructions. These notes were standard instructions on broad topics and opioid mentions within them that did not indicate that the patient in the encounter had taken or been prescribed an opioid. The most effective exclusion criteria were either a heading of "patient education" in the note text or the classification of CCD notes labeled as patient education. Not all patient education notes were explicitly marked in this way, but, when they were, they were excluded.

After the above were excluded, to detect opioid terms, keyword lists were used (referenced in Code component phase 1: Initial code and search term lists). A significant difficulty in finding terms is that they are frequently misspelled. The initial lists included many misspellings, but not all misspellings can be foreseen. Therefore, search capability was enhanced by implementing spelling correction based on named entity recognition (NER). NER is an NLP approach, generally performed with machine learning models, that attempts to find categories of terms by learning what those categories look like. Commonly learned categories are "named entities," alternatively called "proper nouns." The use case is most common for these classes of words because there will never be a definitive list of, for example, all the first names of people in the world. However, a machine may be trained to recognize when a word is someone's first name. The classes need not be proper nouns, though, and in this case, the class added to a base NER system from the NLP package spaCy was "drug term" (22). Annotated examples of data were provided for training, where each training example was a sentence or portion of text with the drug terms identified. Based on these data, the new NER was now able to recognize drug terms as a category in addition to the other kinds of categories it already knew (names, organizations, dates, etc.).

After the NER was built, it was employed in conjunction with the keyword lists. Terms identified by the NER that were exact matches to the opioid list were considered definitive opioid mentions. Terms that were exact matches to nonopioid lists were excluded. If the term was on neither list, the NLP processor attempted to map the term to its correct spelling on the opioid list using Jaro and Levenshtein's string similarity metrics (23,24). If the Jaro string similarity (0 to 1 scale) was greater than 0.90, or the Jaro similarity greater than 0.85 and Levenshtein edit distance equal to 1, based on what was initially seen in early stages of development, the candidate term was proposed as a match to that opioid. If there were multiple matches, the most similar one was considered to be the match. Each unique pair of proposed spelling corrections and original spellings was saved, along with a record of which encounters corrections were proposed for. These were later reviewed by the clinicians on the team to determine if the term was a correct spelling correction or an incorrect one. The encounters that contained the correct spelling corrections were then updated to reflect the presence of that opioid in that encounter.

After the opioid term was found through the means described above, the next step was to determine if the term should be classified as a negated term. If the term was found in the medication data with a "status," that specific term mention would be ruled out if the status was defined as "deleted." Outside of medication data, a negation cue detector was used to determine if that term was negated ("did not use term") (25). If the term was negated, it was removed from the list of matched opioids.

Within the same span of text over which negation detection was performed—roughly, a sentence or a row in a medications table—the NER was also used to detect dates. No additional training for date detection was needed for drug entity recognition because the base model already included the ability to recognize dates. If a date was detected and a regular expression could detect the year of the date, and that year was after 2016, the term was also excluded.

After the rule-outs for negation and dates had been performed, a list of opioids involved for the encounter remained. Each of these terms then had to be mapped to one of the 17 drug categories. Using the DMI principal variants, which are mapped to the keywords (dilaudid mapped to hydromorphone), and with the aid of the clinicians on the project, every term was either mapped to one of the opioid-involved or opioid antagonist categories. Opioids that did not fall in the mentioned opioid categories were captured and mapped as "other opioids." Generic opioid mentions, terms like "opioid," "opiate," or "narcotics," were mapped as "unspecified opioids."

Identifying opioid overdoses

The goal of the NLP component was to identify mentions of acute opioid overdose using the following processes:

- Perform up-front exclusions;
- Find overdose terms;
- Detect rule-outs (negated terms like "did not use" and dates beyond the scope of the survey [before or after 2016]); and
- Find nearby mentioned drugs to see if they included opioids.

Up-front exclusions were conducted in the same way as described above for opioid involvement and specific opioid mentions. To find overdose terms, the NLP processor looked for a mention of "overdose" or "poisoning." This was termed an "overdose mention." An overdose mention included generic and brand name drug mentions of antagonists and misspellings of "overdose." Overdose misspellings were determined up front, instead of during run time for the opioid misspellings. Using a semantic model of the data built earlier, which contained all the words in the data set, the similarity of every word's spelling in the vocabulary to "overdose" was calculated and the top 50 most similarly spelled words were kept. Among those, it was manually determined by clinicians and NLP analysts which misspellings to include for overdose.

Determining if the opioid overdose mention was negated was similar to the method used for opioid involvement as described above. Additionally, date detection using NER was also implemented. If a year could be determined and that year was before or after 2016, the mention was excluded.

After a nonnegated overdose mention was found, a search was then made for nearby drugs. First, the sentence containing the overdose mention was searched. If that sentence contained an opioid keyword, the encounter was classified as an opioid overdose. If not, the sentence was then searched for nonopioid drug terms. If a match was made there, that overdose mention was considered to not be an opioid overdose. If neither of those matches occurred, the previous and following sentences were searched for an opioid term and, if a match was found, the mention was classified as an opioid overdose. If no match was found, the mention was considered to not be an opioid overdose.

Validating and refining NLP processors against gold standard annotated data

The gold standard annotation data set was used to develop the NLP component of the Enhanced Opioid Identification Algorithm. The enhanced algorithm seeks to identify opioid-involved encounters by determining opioid use, the opioid mentioned (13 opioids and 2 antagonists), and if the encounter had evidence of an opioid overdose.

The portions of the annotation data set were used to develop or test the accuracy of the NLP processors. Annotators were asked to identify evidence of the opioid-involved encounters, illicit or prescription opioids taken by the patient, and the nature of the patient's opioid use. The annotators were instructed to document the exact verbiage in the clinical notes of mentioned opioids according to their opioid category, including the 13 categories in the algorithm's output. The annotators also identified if the patient received a diagnosis code related to opioid use or overdose.

The gold standard annotation data set was partitioned into a set to inform the development of the NLP processors (the development set) and a set to evaluate the performance of the NLP processors (the evaluation set). The development set for the opioid involvement question was the 50 encounters that were annotated by all three annotators and whose agreement was evaluated. The development set for the opioid overdose question included the same 50 encounters plus 7 additional encounters positively identified as involving an opioid overdose. In both cases, the evaluation sets were the remainder of the data set (those not included in development). The performance of the algorithms was measured against those evaluation sets and is reported in Tables 5–10.

Performance based on results obtained from the code component of the Enhanced Opioid Identification Algorithm alone is reported in Tables 5 and 6. Performance based on results obtained from the NLP component is reported in Tables 7 and 8. Performance based on results obtained by the full algorithm, including both the code and NLP components, is reported in Tables 9 and 10. Only the top-level yes or no question for identifying opioid involvement was evaluated (Is there any evidence of opioid use by the patient?*)*; results do not include category-specific evaluation. "Recall" (also known as sensitivity) is the percentage of true positives over the sum of true positives and false negatives. "Precision" (also known as positive predictive value) is the percentage of true positives over the sum of true positives and false positives. F1 is the harmonic mean of recall and precision, a common measure of algorithm performance. MCC is Matthews correlation coefficient (identical to Pearson's phi coefficient), which provides a measure balanced over true and false negatives and positives. All calculations are based on numbers found in the confusion matrix, where:

- The cell for annotator positive and algorithm component positive equals true positives;
- The cell for annotator positive and algorithm component negative equals false negatives;
- The cell for annotator negative and algorithm component positive equals false positives; and
- The cell for annotator negative and algorithm component negatives equals true negatives.

In interpreting these results, an important point to remember is a false negative for either component of the enhanced algorithm (NLP or code) may be the result of information not being present in the text or code-based diagnoses, respectively, as opposed to the information being present but the algorithm's failing to identify it. Additionally, one final point to remember is that the assumption of ground truth is annotator data, and it is possible that information was incorrectly classified during the annotation due to human error.

Results

Annotation Results

- The code component of the Enhanced Opioid Identification Algorithm correctly identified 317 opioid-involved encounters that the annotators also identified as opioid-involved in the gold standard annotation data set (Table 5). The code component also identified 630 nonopioid-involved encounters that were categorized as nonopioid-involved by the annotators. However, 926 encounters were incorrectly identified as nonopioid-involved by the code component (false negatives) and 10 encounters were incorrectly identified as opioid-involved by the coded component of the enhanced algorithm (false positives). Overall, there was a positive correlation (0.3) between opioid-involved encounters identified by annotators in the gold standard annotation data set and those identified by the code component (Table 6). This result was anticipated because the clinical notes often mention opioid drugs, particularly therapeutics, for which there is no corresponding medical code in any of the coded fields.
- The code component correctly identified 58 overdose encounters (true positives) and 1,816 nonoverdose encounters (true negatives) (Table 5). The code component incorrectly identified one encounter as an opioid overdose (false positive) and incorrectly identified eight encounters

as nonopioid overdose encounters (false negatives). Overall, there was a positive correlation (0.93) between opioid overdose encounters identified by annotators in the gold standard data set and overdose encounters identified by the code component (Table 6).

- The NLP component of the Enhanced Opioid Identification Algorithm and the clinical annotators identified 1,178 opioid-involved encounters (true positives) and 487 encounters as nonopioid-involved (true negatives) in the gold standard data set (Table 7). The NLP component of the enhanced algorithm incorrectly identified 153 encounters as opioid-involved (false positives) and incorrectly identified 65 encounters as nonopioid-involved (false negatives). Overall, there was a positive correlation (0.74) between the opioid-involved encounters identified by the NLP component of the algorithm and the annotator's gold standard data set (Table 8).
- For opioid overdoses, the NLP component algorithm correctly identified 55 encounters as opioid overdoses (true positives) and 1,809 encounters as nonopioid overdoses (true negatives) (Table 7). The NLP component incorrectly identified 8 encounters as opioid overdoses (false positives) and incorrectly identified 11 encounters as nonopioid overdoses (false negatives). There was a positive correlation of 0.85 between the opioid overdoses identified by the NLP component and the gold standard annotation data set (Table 8).
- The Enhanced Opioid Identification Algorithm (including both the coded and NLP components) correctly identified 1,204 opioid-involved encounters (true positives) and 485 nonopioid-involved encounters (true negatives) (Table 9). The algorithm incorrectly identified 155 encounters as opioid-involved (false positives) and 39 encounters as nonopioid-involved (false negatives). There was a positive correlation (0.77) of opioid-involved encounters identified by the enhanced algorithm and gold standard annotation data set (Table 10).
- The Enhanced Opioid Identification Algorithm correctly identified 64 opioid overdoses (true positives) and 1,808 nonopioid overdose encounters (true negatives). The algorithm incorrectly identified nine encounters as opioid overdose encounters (false positives) and two encounters as nonopioid overdose encounters (false negatives) (Table 9). Overall, there was a positive correlation (0.92) between opioid overdose encounters identified by the enhanced algorithm and the gold standard annotation data set (Table 10).

The code component of the algorithm only identified 25.5% of the annotator-identified opioid-involved encounters (Table 6), while the NLP component alone performed better at identifying annotator-identified opioid-involved encounters, with a recall of 94.8% (Table 8). However, the most accurate version of the algorithm was when the coded and NLP component were both used together to identify opioid-involved encounters. Table 10 reports a recall percentage of 96.9% when the code and NLP components were used to identify opioid-involved encounters.

Results of the Enhanced Algorithm in the 2016 NHCS

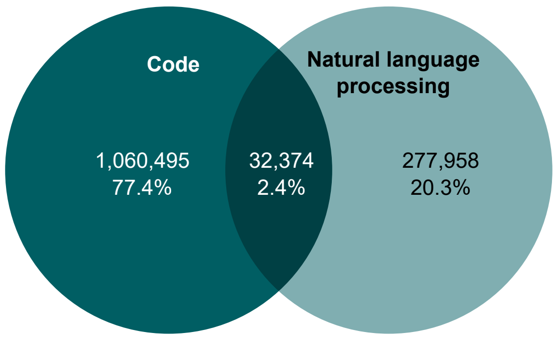
The Enhanced Opioid Identification Algorithm using the code and NLP components identified 1,370,827 opioid-involved and 21,693 opioid overdose encounters in the 2016 NHCS. By comparison, if the enhanced algorithm relied on ICD–10–CM diagnosis codes exclusively to identify opioid-involved hospital encounters, only 112,534 ED and 99,486 inpatient encounters would be identified. Table 11 shows the percentage of opioid-involved encounters identified in the 2016 NHCS ED and inpatient department. The percentage of 2016 NHCS encounters that were identified as opioid overdoses in the ED and inpatient department are shown in Table 12. Table 13 reports the percentage of ED and inpatient encounters in the 2016 NCHS in the 17 drug categories of interest within the total encounters with at least one drug mention.

- 11.5% of ED and 21.8% of inpatient encounters were identified as opioid-involved by the Enhanced Opioid Identification Algorithm (Table 11).
- 0.2% of ED and 0.3% of inpatient encounters were identified as opioid overdoses by the Enhanced Opioid Identification Algorithm (Table 12).
- The most identified opioid mention in ED encounters was morphine (34.3%) and in inpatient encounters was oxycodone (37.9%) (Table 13).

Figure 1 shows the number of total opioid-involved encounters detected by the code component, the NLP component, and the overlap between the two methodologies. Of the total opioid-involved encounters, 20.3% were identified by the NLP component only. This shows the value of using both the coded and NLP components for identifying opioid-involved encounters, because the NLP component was able to identify 277,958 opioid-involved encounters that would not have been identified if only using codes. Figure 2 reports on the percentage of opioid overdose encounters identified by both components and the overlap between the code component and the NLP component. The NLP component identified 2.9% of opioid overdose encounters that were not identified by the code component of the algorithm, demonstrating that compared with its use for identification of opioid-involved encounters, the NLP component of the algorithm was less successful at identifying opioid overdose encounters not identified by the code component of the algorithm.

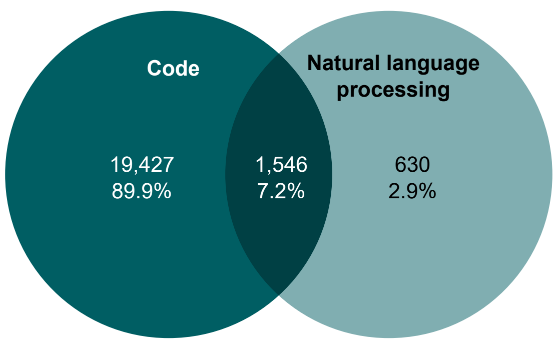
Findings were similar for identifying opioid-involved encounters in the ED (Table 14) and inpatient settings (Table 15), although the NLP component identified a greater percentage of opioid-involved ED encounters (24.7%) compared with opioid-involved inpatient encounters (13.9%). For opioid overdose encounters in both settings,

Figure 1. Total number of opioid-involved encounters identified by the code, natural language processing, and both components of the Enhanced Opioid Identification Algorithm



SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Figure 2. Total number of opioid overdose encounters identified by the code, natural language processing, and both components of the Enhanced Opioid Identification Algorithm



SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

approximately 90% were identified with only the code component.

Analytic Considerations and Limitations

The Enhanced Opioid Identification Data Set was created using 2016 NHCS data, which is not nationally representative and cannot be used to make national estimates. In addition, each of the data sources for the 2016 NHCS (Vizient, EHR, and UB–04 administrative claims) have limitations that affected how each component identified opioid-involved encounters. The sections below describe the limitations of the methodologies and data that must be considered when interpreting the results of the enhanced algorithm.

Limitations of the Code Component of the Enhanced Opioid Identification Algorithm

Limitations of identifying opioid-involved encounters and specific opioids mentioned

The code component identified opioid-involved encounters from medical codes such as diagnosis, procedure, and medication codes. While diagnosis and service codes provide a standardized and efficient way to identify opioid-involved encounters, there are limitations associated with this method. Several hospitals submitted data with incomplete or nonstandardized diagnosis, procedure, and medication information. In addition to searching for codes, the code component of the algorithm included a keyword search of the medical code labels. Unlike the NLP component, the code-based component of the Enhanced Opioid Identification Algorithm's keyword searches assumed these labels were standardized (not truncated or misspelled). Therefore, truncated and misspelled words in code labels would not match included search terms. Encounters that were missing both codes and labels would also have been excluded from the code and keyword searches.

The search for diagnostic medical codes relied on the use of ICD–10–CM codes. However, the number of available ICD–10–CM codes in the 2016 NHCS varied by data source. UB–04 administrative claims and Vizient files were limited to a maximum of 28 diagnostic codes per encounter, including up to 3 reason for visit fields and up to 25 diagnosis fields. However, EHR data provided an unlimited number of diagnostic codes per encounter. In the EHR data, some diagnoses were embedded within the clinical notes and had to be extracted from the text to be searched by the code component of the algorithm. It is possible that some embedded diagnostic information was not extracted. Across all data sources, diagnostic information provided in alternative coding systems were mapped to their equivalent ICD–10–CM codes. Diagnosis codes that could not be mapped to ICD–10–CM were not searched by the code component of the Enhanced Opioid Identification Algorithm.

The code component also used service codes to identify opioid involvement and specific types of opioids mentioned. Some hospitals did not provide procedure codes, but if a label was present that described a procedure or service, a keyword search was used to identify opioid-involved encounters. Labels that contained erroneous medical terms or drugs (catheter) or terms that could potentially be misclassified as a brand name opioid (griseofulvin [an antifungal]) were not searched by the code component of the algorithm. Service labels that contained negation terms (opioid contraindicated) were not searched by the code component of the algorithm.

Opioid-involved drug screens or confirmatory tests were identified using the LOINC laboratory code system. However, the presence of an opioid-involved laboratory code was not strong enough evidence to classify an encounter as opioid-involved. The presence of a laboratory code identifies whether a drug screen was ordered during the hospital encounter, but it does not mean that the patient took the test or that the results of the screening would be reported in the data. If a hospital did not have a laboratory code but provided a label, a keyword search was used to identify the opioid-involved laboratory tests.

The quality and quantity of the medication data varied by source in the 2016 NHCS. Medication data were not available for UB–04 administrative claims data. Among other sources, the medication data provided had some codes, some labels, but rarely had both a code and a label. When medication codes were available, the enhanced algorithm identified opioid-involved medications using the RxNorm or SNOMED–CT coding systems.

Limitations of identifying opioid overdose encounters

Opioid overdoses were identified in the coded algorithm using diagnosis codes. First-listed diagnosis could not be used to identify if an opioid overdose was the reason for a hospital encounter because some encounters had missing or multiple primary diagnoses. There were also issues with establishing temporality of an overdose encounter. Encounter dates could not be used reliably to identify when an opioid overdose occurred or if it occurred during the hospital encounter. For the algorithm, an opioid overdose was included if the diagnosis code had a missing date or if the diagnosis code had a date within calendar year 2016.

Limitations of the NLP Component of the Enhanced Opioid Identification Algorithm

Limitations of identifying opioid-involved encounters

The NLP component of the Enhanced Opioid Identification Algorithm analyzed clinical notes from ED and inpatient encounters. The NLP is limited to available clinical notes data; encounters without clinical notes were not included in the NLP component of the algorithm. Of the 45 million total encounters in the 2016 NHCS data set, only 11.1% had clinical notes. EHR (CCD and Custom Extract) hospitals were the only hospitals that submitted clinical notes. UB–04 administrative claims and Vizient did not submit clinical notes and were not searched by the NLP component of the enhanced algorithm.

The formatting of clinical notes can affect the performance of the NLP component of the Enhanced Opioid Identification Algorithm. Clinical notes were not uniform in data quality. In the case of one hospital, all note texts were limited to the first 256 characters, likely omitting much of the note. Some notes lacked punctuation, which can inhibit the algorithm's effectiveness in identifying dates or negations. In some cases, notes were only phrases or fragments of sentences (nurse at bedside), and opioid involvement was not indicated in the text.

Nearly all NLP methods employed are inexact, as textual representations of information vary widely. It is uncommon for all ways of encoding information in natural language to be captured by any algorithm. NLP negation exclusions can erroneously exclude opioid-involved encounters as false negatives and identify nonopioid-involved encounters as false positives.

NER did not find all drug mentions and sometimes found terms that were not drugs but were similar to them (vitamins and minerals). This limitation is mitigated through automatic and manual filtering of terms that are not relevant to the case definition. Spelling correction could incorrectly map an NER-identified candidate drug term to an opioid spelling or could fail to do the appropriate mapping. The latter limitation was resolved while the former was considered a more serious error, which is why the "human-in-the-loop" approach was used by project clinicians vetting suggestions.

Limitations of identifying opioid overdose encounters

There are limitations to the NLP component of the enhanced algorithm that identified opioid overdoses. Date exclusions did not find all relevant dates, which may have undercounted the number of opioid overdoses in the clinical notes. If a date was not in the same sentence as a targeted opioid, it was not considered a relevant date and the encounter was not identified as an opioid overdose. Unforeseen textual cues concerning opioid overdoses could go undetected. If the opioid associated with the overdose was not in the same, previous, or following sentence as the overdose trigger term (overdose or poisoning), the overdose would be undetected by the NLP algorithm. This limitation may be minor, as annotators indicated that overdoses rarely occurred in the data they annotated without the presence of an explicit mention of overdose or poisoning.

Discussion

The methodology detailed in this report seeks to improve the identification of opioid-involved and opioid overdose hospital encounters in NHCS data. Medical codes provide limited information about context in the hospital encounters and can miss opioid-involved encounters that are embedded in free text data, like hospital clinical notes. The Enhanced Opioid Identification Algorithm developed in this FY18 PCORTF project leverages both code-based and textual elements of NHCS data to identify opioid-involved encounters that would otherwise be overlooked by searching exclusively for medical codes.

By setting, 22% of inpatient department encounters and 12% of ED encounters were identified as opioid-involved. The inpatient setting had a higher prevalence of encounters due to the inclusion of opioid-involved medical procedures, medications, and services that are often a part of patient care. Earlier iterations of opioid-identification algorithms identified a much lower percentage of opioid-involved ED encounters in the 2013 NHCS (0.8%) (3). This disparity likely reflects several factors, particularly the use of an expanded set of case definition inclusion criteria for the PCORTF projects. The Enhanced Opioid Identification Algorithm was designed to capture all forms of past and present opioid use, including use of prescription opioids as directed by a physician, misuse of prescription opioids in a manner other than as directed by a physician, and any use of illicit opioids. In addition, the algorithms also flag encounters indicating that an opioid was prescribed in the past (medication history), given during the encounter, or prescribed upon discharge. Future iterations of the algorithm could describe the nature of opioid-involvement in hospital encounters, such as identifying if an opioid prescription or service was given during the encounter.

The performance of the Enhanced Opioid Identification Algorithm against the gold standard data set annotated by clinicians highlighted the limitations of using medical codes to identify opioid-involved encounters. The code component of the enhanced algorithm had more false negatives compared with the NLP component. The coded component falsely identified 926 encounters as not opioid-involved compared with 65 encounters falsely identified as not opioid-involved in the NLP component of the enhanced algorithm. As a result, the MCC correlation between the code component of the enhanced algorithm was significantly weaker (0.30) than the NLP component of the enhanced algorithm (0.74). The addition of the NLP component of the enhanced algorithm allowed for the identification of opioid-involved encounters that would not have been found relying on medical codes.

The current methodology of the Enhanced Opioid Identification Algorithm identifies opioid-involved encounters, but it does not identify behaviors associated with opioid use disorders. Future considerations include improving the algorithm's sensitivity to identify opioid use disorders and drug-seeking behavior in hospital encounter data. Additionally, the upcoming FY19 PCORTF project will leverage the methodology used for developing the Enhanced Opioid Identification Algorithm in the development of an algorithm that will identify opioid-involved encounters that have co-occurring substance use disorders and mental illness.

Opioid overdoses were rare in each setting. Opioid overdoses were identified in 0.2% of ED encounters and 0.3% of inpatient encounters. These findings are similar to those found in other syndromic surveillance systems during the same time period. The CDC's National Syndromic Surveillance Program and ESOOS program reported approximately 0.2% of ED encounters between July 2016 and September 2017 were suspected opioid overdoses (26). A systematic review of 13 studies of hospital admissions found between 0.06% and 2.5% of inpatient encounters were due to opioid overdose (27). The code component of the enhanced algorithm identified almost all of the overdose encounters in both hospital settings. The code component of the enhanced algorithm identified opioid overdoses exclusively using ICD–10–CM diagnosis codes. In the future, it may be helpful to incorporate symptom-based and procedure-based indicators to predict the likelihood of an overdose event occurring in present or future hospital encounters. Future iterations of the algorithm could expand the definition of an opioid overdose to include patients who had an overdose before the survey year.

The Enhanced Opioid Identification Algorithm can improve and optimize the identification of opioid-involved encounters and overdoses in hospital data by revealing opioid-involved encounters and overdoses, identified from clinical notes, that would not have been identified just by relying on medical codes. Further, the Enhanced Opioid Identification Data Set contains linkage variables, allowing researchers the ability to examine more details about opioid-involved hospital encounters using the 2016 NHCS data set or identify mortality using the 2016–2017 NDI and DIM files. After NHCS becomes nationally representative, the enhanced algorithm can be a powerful tool to identify national trends in the opioid epidemic.

References

1. Spencer MR, Flagg LA, Jackson G, DeFrances C, Hedegaard H. National Hospital Care Survey demonstration projects: Opioid-involved emergency department visits, hospitalizations, and deaths. National Health Statistics Reports; no 141. Hyattsville, MD: National Center for Health Statistics. 2020.

2. Brown AM, DeFrances C, Crane E, Cai R, Naeger S. Identification of substance-involved emergency department visits using data from the National Hospital Care Survey. National Health Statistics Reports; no 114. Hyattsville, MD: National Center for Health Statistics. 2018.

3. National Center for Health Statistics. National Hospital Care Survey. Available from: https://www.cdc.gov/nchs/nhcs/index.htm.

4. Centers for Disease Control and Prevention. Opioid basics: Commonly used terms. 2020. Available from: https://www.cdc.gov/drugoverdose/opioids/terms.html.

5. Centers for Disease Control and Prevention. International Classification of Diseases, 10th Revision, Clinical Modification (ICD–10–CM). 2020. Available from: https://www.cdc.gov/nchs/icd/icd10cm.htm.

6. Jhung MA, Budnitz DS, Mendelsohn AB, Weidenbach KN, Nelson TD, Pollock DA. Evaluation and overview of the National Electronic Injury Surveillance

System–Cooperative Adverse Drug Event Surveillance Project (NEISS–CADES). Med Care 45(10 Supl 2): S96–S102. 2007.

7. Cash RE, Kinsman J, Crowe RP, Rivard MK, Faul M, Panchal AR. Naloxone administration frequency during emergency medical service events—United States, 2012–2016. MMWR Morb Mortal Wkly Rep 67(31):850–3. 2018. DOI: 10.15585/mmwr.mm6731a2.

8. Knowlton A, Weir BW, Hazzard F, Olsen Y, McWilliams J, Fields J, Gaasch W. EMS runs for suspected opioid overdose: Implications for surveillance and prevention. Prehosp Emerg Care 17(3):317–29. 2013. DOI: 10.3109/10903127.2013.792888.

9. Grover JM, Alabdrabalnabi T, Patel MD, Bachman MW, Platts-Mills TF, Cabanas JG, Williams JG. Measuring a crisis: Questioning the use of naloxone administrations as a marker for opioid overdoses in a large U.S. EMS system. Prehosp Emerg Care 22(3):281–9. 2018.

10. Monitto CL, Kost-Byerly S, White E, Lee CKK, Rudek MA, Thompson C, Yaster M. The optimal dose of prophylactic intravenous naloxone in ameliorating opioid-induced side effects in children receiving intravenous patient-controlled analgesia morphine for moderate to severe pain: A dose finding study. Anesth Analg 113(4):834–42. 2011.

11. Chidambaran V, Olbrecht V, Hossain M, Sadhasivam S, Rose J, Meyer MJ. Risk predictors of opioid-induced critical respiratory events in children: Naloxone use as a quality measure of opioid safety. Pain Med 15(12):2139–49. 2014.

12. Gkegkes ID, Minis EE, Iavazzo C. Oxycodone/naloxone in postoperative pain management of surgical patients. J Opioid Manag 14(1):52–60. 2018.

13. Vivolo-Kantor AM, Seth P, Gladden RM, Mattson CL, Baldwin GT, Kite-Powell A, Coletta MA. Vital Signs: Trends in emergency department visits for suspected opioid overdoses—United States, July 2016–September 2017. MMWR Morb Mortal Wkly Rep 67(9):279–85. 2018. DOI: 10.15585/mmwr.mm6709e1.

14. Drug Enforcement Agency. Slang terms and code words: A reference for law enforcement personnel. DEA-HOU-DIR-022-18. 2018.

15. Ward PJ, Rock PJ, Slavova S, Young AM, Bunn TL, Kavuluru R. Enhancing timeliness of drug overdose mortality surveillance: A machine learning approach. PloS One 14(10):e0223318. 2019.

16. Clinical Pharmacology. Elsevier. Available from: https://www.clinicalpharmacology.com/.

17. American Society of Health-System Pharmacists. AHFS pharmacologic-therapeutic classification system. Available from: https://www.ahfsdruginformation.com/ahfs-pharmacologic-therapeutic-classification/.

18. U.S. National Library of Medicine. Value set authority center. Bethesda, MD. Available from: https://vsac.nlm.nih.gov/.

19. Hazlehurst B, Green CA, Perrin NA, Brandes J, Carrell DS, Baer A, et al. Using natural language processing of clinical text to enhance identification of opioid-related overdoses in electronic health records data. Pharmacoepidemiol Drug Saf 28(8):1143–51. 2019. DOI: 10.1002/pds.4810.

20. Lingeman JM, Wang P, Becker W, Yu H. Detecting opioid-related aberrant behavior using natural language processing. AMIA Annu Symp Proc 2017:1179–85. 2018.

21. Artstein R. Inter-annotator agreement. In: Ide N, Pustejovsky J, editors. Handbook of linguistic annotation. Dordrecht, NL: Springer. 297–313. 2017.

22. SpaCy: Industrial-strength natural language processing (NLP) with Python and Cython. Available from: https://github.com/explosion/spacy-models/releases/tag/en_core_web_md-2.2.5.

23. Jaro MA. Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida. JAMA 84(406):414–20. 1989.

24. Levenshtein VI. Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics—Doklady 10(8):707–10. 1966.

25. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 34(5):301–10. DOI: 10.1006/jbin.2001.1029.

26. Vivolo-Kantor AM, Seth P, Gladden RM, Mattson CL, Baldwin GT, Kite-Powell A, Coletta MA. Vital Signs: Trends in emergency department visits for suspected opioid overdoses—United States, July 2016–September 2017. MMWR Morb Mortal Wkly Rep 67(9):279–85. DOI: 10.15585/mmwr.mm6709e1.

27. Danovitch I, Vanle B, Van Groningen N, Ishak W, Nuckols T. Opioid overdose in the hospital setting: A systematic review. J Addict Med 14(1):39–47. 2020.

Table 1. Encounters with at least one *International Classification of Diseases, 10th Revision, Clinical Modification* code in the 2016 National Hospital Care Survey

International Classification of Diseases, 10th Revision, Clinical Modification code	Emergency department		Inpatient	
	Number	Percent	Number	Percent
One or more	6,988,635	99.4	2,556,606	98.6
None	43,669	0.6	35,116	1.4
Total encounters	7,032,304	100.0	2,591,722	100.0

NOTE: Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 2. Encounters with at least one procedure code or label in the 2016 National Hospital Care Survey

Procedure code or label	Emergency department		Inpatient	
	Number	Percent	Number	Percent
One or more	6,320,008	89.9	2,407,076	92.9
None	712,296	10.1	184,646	7.1
Total encounters	7,032,304	100.0	2,591,722	100.0

NOTE: Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 3. Encounters with at least one laboratory code or label in the 2016 National Hospital Care Survey

Laboratory code or label	Emergency department		Inpatient	
	Number	Percent	Number	Percent
One or more	505,693	7.2	407,966	15.7
None	6,526,611	92.8	2,183,756	84.3
Total encounters	7,032,304	100.0	2,591,722	100.0

NOTE: Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 4. Encounters with at least one medication code or label in the 2016 National Hospital Care Survey

Medication code or label	Emergency department		Inpatient	
	Number	Percent	Number	Percent
One or more	1,801,395	25.6	972,784	37.5
None	5,230,909	74.4	1,618,938	62.5
Total encounters	7,032,304	100.0	2,591,722	100.0

NOTE: Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 5. Agreement counts between the code component of the Enhanced Opioid Identification Algorithm and the annotated data set

Characteristic	Opioid-involved encounters		Opioid overdose encounters	
	Annotator positive	Annotator negative	Annotator positive	Annotator negative
Algorithm positive	317	10	58	1
Algorithm negative	926	630	8	1,816

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 6. Performance measures of the code component of the Enhanced Opioid Identification Algorithm and the annotated dataset

Characteristic	Opioid involvement	Opioid overdose
Recall[1]	25.5	87.9
Precision[2]	96.9	98.3
F1[3]	20.4	92.8
Matthews correlation coefficient	0.30	0.93

[1]Percentage of correctly identified positives out of all true positives, also known as sensitivity.
[2]Percentage of identified positives that are true positives.
[3]Harmonic mean of recall and precision, a common measure of algorithm performance.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 7. Agreement counts between the natural language processing component of the Enhanced Opioid Identification Algorithm and the annotated data set

Characteristic	Opioid-involved encounters		Opioid overdose encounters	
	Annotator positive	Annotator negative	Annotator positive	Annotator negative
Algorithm positive	1,178	153	55	8
Algorithm negative	65	487	11	1,809

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 8. Performance measures of the natural language processing component of the Enhanced Opioid Identification Algorithm and the annotated data set

Characteristic	Opioid involvement	Opioid overdose
Recall[1]	94.8	83.3
Precision[2]	88.5	87.3
F1[3]	91.5	85.3
Matthews correlation coefficient	0.74	0.85

[1]Percentage of correctly identified positives out of all true positives, also known as sensitivity.
[2]Percentage of identified positives that are true positives.
[3]Harmonic mean of recall and precision, a common measure of algorithm performance.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 9. Agreement counts between the full Enhanced Opioid Identification Algorithm and the annotated data set

Characteristic	Opioid-involved encounters		Opioid overdose encounters	
	Annotator positive	Annotator negative	Annotator positive	Annotator negative
Algorithm positive	1,204	155	64	9
Algorithm negative	39	485	2	1,808

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 10. Performance measures of the full Enhanced Opioid Identification Algorithm and the annotated dataset

Characteristic	Opioid involvement	Opioid overdose
Recall[1]	96.9	97.0
Precision[2]	88.6	87.7
F1[3]	92.5	92.1
Matthews correlation coefficient	0.77	0.92

[1]Percentage of correctly identified positives out of all true positives, also known as sensitivity.
[2]Percentage of identified positives that are true positives.
[3]Harmonic mean of recall and precision, a common measure of algorithm performance.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 11. Number and percent distribution of opioid-involved emergency department and inpatient encounters

Characteristic	Emergency department		Inpatient	
	Number	Percent	Number	Percent
Opioid-involved	805,456	11.5	565,371	21.8
Not opioid-involved	6,226,848	88.5	2,026,351	78.2
Total encounters	7,032,304	100.0	2,591,722	100.0

NOTE: Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 12. Number and percent distribution of opioid overdose emergency department and inpatient encounters

Characteristic	Emergency department		Inpatient	
	Number	Percent	Number	Percent
Opioid-involved	14,728	0.2	6,875	0.3
Not opioid-involved	7,017,576	99.8	2,584,847	99.7
Total encounters	7,032,304	100.0	2,591,722	100.0

NOTE: Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 13. Encounters in the 17 drug categories of interest for the encounters with at least one drug mention

Drug category[1]	Emergency department		Inpatient	
	Number	Percent	Number	Percent
Buprenorphine	2,086	0.3	2,507	0.4
Codeine	49,021	6.1	29,875	5.3
Fentanyl	86,474	10.7	28,301	5.0
Heroin	8,098	1.0	2,704	0.5
Hydrocodone	136,817	17.0	93,590	16.6
Hydromorphone	172,066	21.4	163,979	29.0
Levorphanol	1,729	0.2	235	0.0
Meperidine	6,040	0.7	9,263	1.6
Methadone	12,425	1.5	12,944	2.3
Morphine	276,051	34.3	167,217	29.6
Oxycodone	82,064	10.2	214,528	37.9
Oxymorphone	312	0.0	653	0.1
Tramadol	58,821	7.3	58,953	10.4
Unspecified opioid	26,105	3.2	13,154	2.3
Other opioid	8,916	1.1	6,839	1.2
Naloxone[2]	4,638	0.6	9,627	1.7
Naltrexone[2]	254	0.0	173	0.0

0.0 Quantity more than 0 but less than 0.05.
[1]Drug categories are not mutually exclusive.
[2]Opioid antagonists.

NOTES: This table is based on the 805,456 emergency department and 565,371 inpatient encounters with at least one drug mention. Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 14. Number and percent distribution of emergency department opioid-involved and opioid overdose encounters, by selection method

Selection method	Opioid-involved encounters		Opioid overdose encounters	
	Number	Percent	Number	Percent
Code component only	584,550	72.6	13,234	89.9
Natural language processing component only	199,110	24.7	516	3.5
Code and natural language processing components	21,796	2.7	978	6.6
Total	805,456	100.0	14,728	100.0

NOTE: Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table 15. Number and percent distribution of inpatient department opioid-involved and opioid overdose encounters, by selection method

Selection method	Opioid-involved encounters		Opioid overdose encounters	
	Number	Percent	Number	Percent
Code component only	475,945	84.2	6,193	90.0
Natural language processing component only	78,848	13.9	114	1.7
Code and natural language processing components	10,578	1.9	568	8.3
Total	565,371	100.0	6,875	100.0

NOTE: Data are unweighted and not nationally representative.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Appendix. Supporting Tables

Table I. *International Classification of Diseases, 10th Revision, Clinical Modification* codes used to identify opioid-involved diagnoses in the enhanced algorithm

Category and code	Description
Use	
F11.9–F11.90	Opioid use, unspecified
F11.92–F11.929	Opioid use, unspecified with intoxication
F11.93	Opioid use, unspecified with withdrawal
F11.94	Opioid use, unspecified with opioid-induced mood disorder
F11.95–F11.959	Opioid use, unspecified with opioid-induced psychotic disorder
F11.98–F11.988	Opioid use, unspecified with other specified opioid-induced disorder
F11.99	Opioid use, unspecified with unspecified opioid-induced disorder
Abuse	
F11.1–F11.11	Opioid abuse
F11.12–F11.129	Opioid abuse with intoxication
F11.14	Opioid abuse with opioid-induced mood disorder
F11.15–F11.159	Opioid abuse with opioid-induced psychotic disorder
F11.18–F11.188	Opioid abuse with other opioid-induced disorder
F11.19	Opioid abuse with unspecified opioid-induced disorder
Dependence	
F11.2–F11.21	Opioid dependence
F11.22–F11.229	Opioid dependence with intoxication
F11.23	Opioid dependence with withdrawal
F11.24	Opioid dependence with opioid-induced mood disorder
F11.25–F11.259	Opioid dependence with opioid-induced psychotic disorder
F11.28–F11.288	Opioid dependence with other opioid-induced disorder
F11.29	Opioid dependence with unspecified opioid-induced disorder
Poisoning	
T40.0X1–T40.0X1S	Poisoning by opium, accidental (unintentional)
T40.0X2–T40.0X2S	Poisoning by opium, intentional self-harm
T40.0X3–T40.0X3S	Poisoning by opium, assault
T40.0X4–T40.0X4S	Poisoning by opium, undetermined
T40.1X1–T40.1X1S	Poisoning by heroin, accidental (unintentional)
T40.1X2–T40.1X2S	Poisoning by heroin, intentional self-harm
T40.1X3–T40.1X3S	Poisoning by heroin, assault
T40.1X4–T40.1X4S	Poisoning by heroin, undetermined
T40.2X1–T40.2X1S	Poisoning by other opioids, accidental (unintentional)
T40.2X2–T40.2X2S	Poisoning by other opioids, intentional self-harm
Poisoning—Con.	
T40.2X3–T40.2X3S	Poisoning by other opioids, assault
T40.2X4–T40.2X4S	Poisoning by other opioids, undetermined
T40.3X1–T40.3X1S	Poisoning by methadone, accidental (unintentional)
T40.3X2–T40.3X2S	Poisoning by methadone, intentional self-harm
T40.3X3–T40.3X3S	Poisoning by methadone, assault
T40.3X4–T40.3X4S	Poisoning by methadone, undetermined
T40.4X1–T40.4X1S	Poisoning by other synthetic narcotics
T40.4X2–T40.4X2S	Poisoning by other synthetic narcotics, self-harm
T40.4X3–T40.4X3S	Poisoning by other synthetic narcotics, assault
T40.4X4–T40.4X4S	Poisoning by other synthetic narcotics, undetermined
T40.601–T40.601S	Poisoning by unspecified narcotics, accidental
T40.602–T40.602S	Poisoning by unspecified narcotics, intentional self-harm
T40.603–T40.603S	Poisoning by unspecified narcotics, assault
T40.604–T40.604S	Poisoning by unspecified narcotics, undetermined
T40.691–T40.691S	Poisoning by other narcotics, accidental (unintentional)
T40.692–T40.692S	Poisoning by other narcotics, intentional self-harm
T40.693–T40.693S	Poisoning by other narcotics, assault
T40.694–T40.694S	Poisoning by other narcotics, undetermined
Adverse effects	
T40.0X5–T40.0X5S	Adverse effect of opium
T40.2X5–T40.2X5S	Adverse effect of other opioids
T40.3X5–T40.3X5S	Adverse effect of methadone
T40.4X5–T40.4X5S	Adverse effect of other synthetic narcotics
T40.605–T40.605S	Adverse effect of unspecified narcotics
T40.695–T40.695S	Adverse effect of other narcotics
Underdosing	
T40.0X6–T40.0X6S	Underdosing of opium
T40.2X6–T40.2X6S	Underdosing of other opioids
T40.3X6–T40.3X6S	Underdosing of methadone
T40.4X6–T40.4X6S	Underdosing of other synthetic narcotics
T40.606–T40.606S	Underdosing of unspecified narcotics
T40.696–T40.696S	Underdosing of other narcotics
Miscellaneous	
Z79.891	Long term (current) use of opiate analgesic
R78.1	Finding of opiate drug in blood
P04.14	Newborn affected by maternal use of opiates
P96.1	Neonatal withdrawal symptoms from maternal use of drugs of addiction

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Table II. RxNorm, Systemized nomenclature of medicine–Clinical terms, and Healthcare common procedure coding system medication and procedural codes used to identify generic opioid agonists and antagonists in the enhanced algorithm

Opioid type	RxNorm concept unique identifier (RxCUI)	Healthcare common procedure coding system and *International Classification of Diseases, 10th Revision* Procedure Coding System[1]	Systemized nomenclature of medicine–Clinical terms (substance)
Opioid agonists (generic)			
Buprenorphine	1819	J0570, J0571	387173000
Codeine	2670	J0745	387494007
Fentanyl	4337	J1810, J3010	373492002
Heroin	3304	…	387341002
Hydrocodone	5489	…	372671002
Hydromorphone	3423	J1170	44508008
Levorphanol	6378	J1960	387275004
Meperidine	6754	J2175, J2180	387298007
Methadone	6813	J1230, HZ81ZZZ, HZ91ZZZ	387286002
Morphine	7052	J2270, J2274	373529000
Oxycodone	7804	…	55452001
Oxymorphone	7814	J2410	24751001
Tramadol	10689	…	386858008
Opioid antagonists (generic)			
Naloxone	7242	J2310, HZ85ZZZ, HZ95ZZZ	372890007
Naltrexone	7243	J2315, HZ84ZZZ, HZ94ZZZ	373546002

… Category not applicable.
[1]Procedure codes related to medication management and pharmacotherapy. For this report, the code-based algorithm was primarily applied to opioid antagonist or agonist name to capture codes.

NOTE: Brand names, drug variations, and additional Systemized nomenclature of medicine–Clinical terms codes can be found in supplemental data for this report.

SOURCE: National Center for Health Statistics, National Hospital Care Survey, 2016.

Vital and Health Statistics Series Descriptions

Active Series

Series 1. Programs and Collection Procedures
Reports describe the programs and data systems of the National Center for Health Statistics, and the data collection and survey methods used. Series 1 reports also include definitions, survey design, estimation, and other material necessary for understanding and analyzing the data.

Series 2. Data Evaluation and Methods Research
Reports present new statistical methodology including experimental tests of new survey methods, studies of vital and health statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. Reports also include comparison of U.S. methodology with those of other countries.

Series 3. Analytical and Epidemiological Studies
Reports present data analyses, epidemiological studies, and descriptive statistics based on national surveys and data systems. As of 2015, Series 3 includes reports that would have previously been published in Series 5, 10–15, and 20–23.

Discontinued Series

Series 4. Documents and Committee Reports
Reports contain findings of major committees concerned with vital and health statistics and documents. The last Series 4 report was published in 2002; these are now included in Series 2 or another appropriate series.

Series 5. International Vital and Health Statistics Reports
Reports present analytical and descriptive comparisons of U.S. vital and health statistics with those of other countries. The last Series 5 report was published in 2003; these are now included in Series 3 or another appropriate series.

Series 6. Cognition and Survey Measurement
Reports use methods of cognitive science to design, evaluate, and test survey instruments. The last Series 6 report was published in 1999; these are now included in Series 2.

Series 10. Data From the National Health Interview Survey
Reports present statistics on illness; accidental injuries; disability; use of hospital, medical, dental, and other services; and other health-related topics. As of 2015, these are included in Series 3.

Series 11. Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey
Reports present 1) estimates of the medically defined prevalence of specific diseases in the United States and the distribution of the population with respect to physical, physiological, and psychological characteristics and 2) analysis of relationships among the various measurements. As of 2015, these are included in Series 3.

Series 12. Data From the Institutionalized Population Surveys
The last Series 12 report was published in 1974; these reports were included in Series 13, and as of 2015 are in Series 3.

Series 13. Data From the National Health Care Survey
Reports present statistics on health resources and use of health care resources based on data collected from health care providers and provider records. As of 2015, these reports are included in Series 3.

Series 14. Data on Health Resources: Manpower and Facilities
The last Series 14 report was published in 1989; these reports were included in Series 13, and are now included in Series 3.

Series 15. Data From Special Surveys
Reports contain statistics on health and health-related topics from surveys that are not a part of the continuing data systems of the National Center for Health Statistics. The last Series 15 report was published in 2002; these reports are now included in Series 3.

Series 16. Compilations of Advance Data From Vital and Health Statistics
The last Series 16 report was published in 1996. All reports are available online; compilations are no longer needed.

Series 20. Data on Mortality
Reports include analyses by cause of death and demographic variables, and geographic and trend analyses. The last Series 20 report was published in 2007; these reports are now included in Series 3.

Series 21. Data on Natality, Marriage, and Divorce
Reports include analyses by health and demographic variables, and geographic and trend analyses. The last Series 21 report was published in 2006; these reports are now included in Series 3.

Series 22. Data From the National Mortality and Natality Surveys
The last Series 22 report was published in 1973. Reports from sample surveys of vital records were included in Series 20 or 21, and are now included in Series 3.

Series 23. Data From the National Survey of Family Growth
Reports contain statistics on factors that affect birth rates, factors affecting the formation and dissolution of families, and behavior related to the risk of HIV and other sexually transmitted diseases. The last Series 23 report was published in 2011; these reports are now included in Series 3.

Series 24. Compilations of Data on Natality, Mortality, Marriage, and Divorce
The last Series 24 report was published in 1996. All reports are available online; compilations are no longer needed.

For answers to questions about this report or for a list of reports published in these series, contact:

Information Dissemination Staff
National Center for Health Statistics
Centers for Disease Control and Prevention
3311 Toledo Road, Room 4551, MS P08
Hyattsville, MD 20782

Tel: 1–800–CDC–INFO (1–800–232–4636)
TTY: 1–888–232–6348
Internet: https://www.cdc.gov/nchs
Online request form: https://www.cdc.gov/info
For e-mail updates on NCHS publication releases, subscribe online at: https://www.cdc.gov/nchs/email-updates.htm.

U.S. DEPARTMENT OF
HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road, Room 4551, MS P08
Hyattsville, MD 20782–2064

For more NCHS Series Reports, visit:
https://www.cdc.gov/nchs/products/series.htm

Series 2, No. 188
CS326147