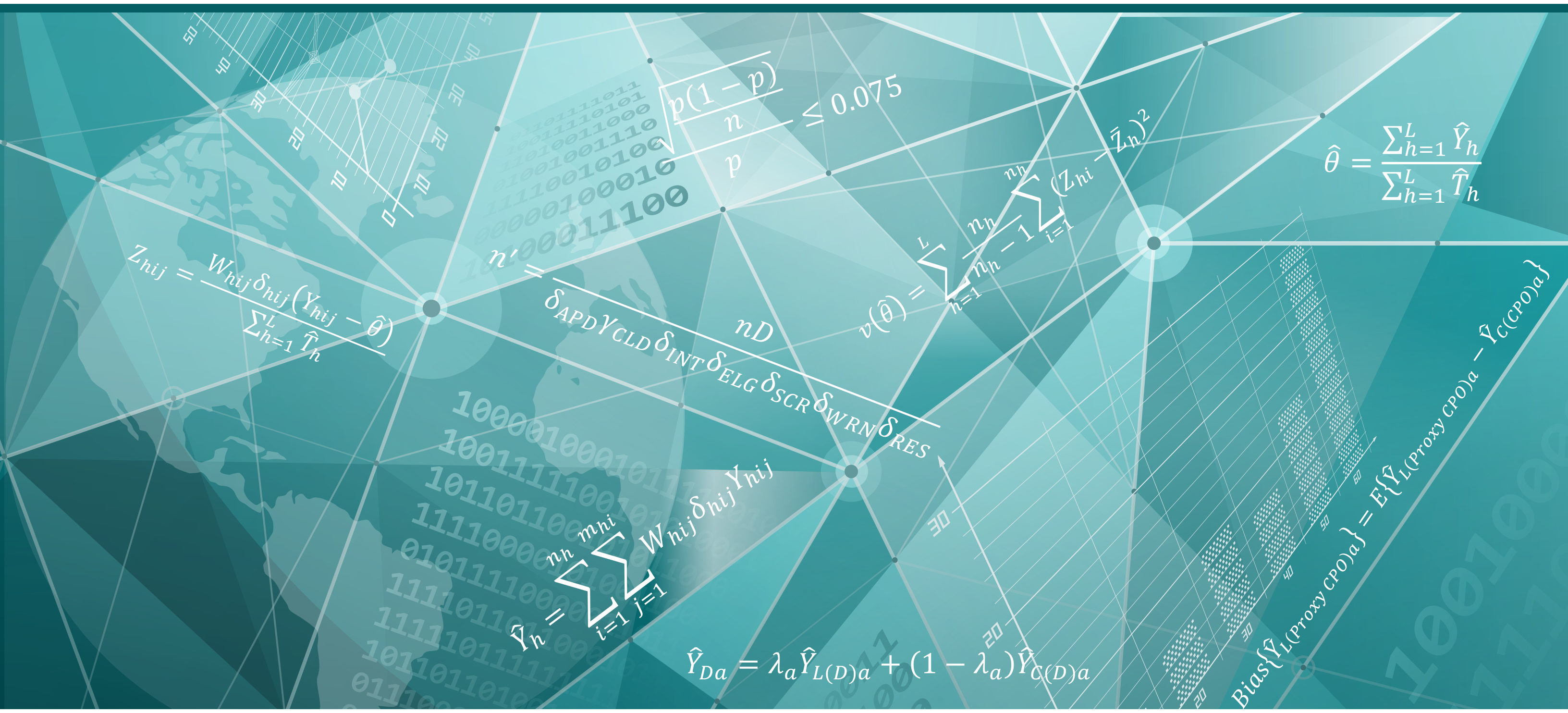
NATIONAL CENTER FOR HEALTH STATISTICS

Vital and Health Statistics

Series 2, Number 193 September 2022



Identifying Co-occurring Disorders Among Patients With an Opioid-involved Hospital Encounter 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

Brown AM, White DG, Adams NB, O'Jiaku-Okorie A, Badwe R, Shaikh S, Adegboye A. Identifying co-occurring disorders among patients with an opioid-involved hospital encounter using National Hospital Care Survey data. National Center for Health Statistics. Vital Health Stat 2(193). 2022. DOI: https://dx.doi.org/10.15620/cdc:119155.

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 193 September 2022

Identifying Co-occurring Disorders Among Patients With an Opioid-involved Hospital Encounter Using National Hospital Care Survey Data

Data and Evaluation Methods Research

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

Hyattsville, Maryland
September 2022

National Center for Health Statistics

Brian C. Moyer, Ph.D., *Director*

Amy M. Branum, Ph.D., *Associate Director for Science*

Division of Health Care Statistics

Carol J. DeFrances, Ph.D., *Acting Director*

Alexander Strashny, Ph.D., *Associate Director for Science*

Contents

Detailed Tables

Identifying Co-occurring Disorders Among Patients With an Opioid-involved Hospital Encounter Using National Hospital Care Survey Data

by Amy M. Brown, M.P.H, Donielle G. White, M.P.H., Nikki B. Adams, Ph.D., 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) Co-occurring Disorders Algorithm, which can be used to identify patients with an opioid-involved hospital encounter who had lifetime diagnoses of both a substance use disorder and a selected mental health issue. Lifetime diagnoses are defined as diagnoses at any point in the past or during the current encounter. This algorithm was created to complement the earlier NHCS Enhanced Opioid Identification Algorithm designed to improve the classification of patients with opioid-involved hospital encounters.

Methods

The Co-occurring Disorders Algorithm incorporates methodology similar to the earlier algorithm, including searches of medical codes and natural language processing (NLP) of free text clinical notes. This algorithm used information about behavioral health conditions that was found in clinical notes to determine whether encounters met all inclusion and exclusion criteria of case definitions. During the development of the algorithm, a subset of NHCS encounters were annotated by reviewers with clinical expertise to identify patients experiencing an opioid-involved hospital encounter with co-occurring disorders. These annotated records allowed for the verification of encounters identified by the algorithm.

Results

In the 2016 NHCS data, the algorithm identified 74,472 opioid-involved emergency department visits and 85,019 opioid-involved inpatient hospital admissions of patients with co-occurring disorders. Approximately 10% of the patients with opioid-involved encounters with a co-occurring disorder were identified using just the NLP component and would not have been identified relying on the code component of the algorithm.

Keywords: opioids • co-occurring disorders • hospitals • natural language processing

Introduction

The National Hospital Care Survey (NHCS), conducted by the National Center for Health Statistics (NCHS), collects data on patient care in hospital-based settings to describe patterns of health care delivery and use in the United States (although currently the data are not nationally representative). In fiscal year 2019, NCHS received funding from the Department of Health and Human Services' Office of the Secretary Patient-Centered Outcomes Research Trust Fund (PCORTF) for a project to develop the Co-occurring Disorders Algorithm. The algorithm would identify lifetime diagnoses of both a substance use disorder (SUD) and a selected mental health issue (MHI) among patients with an opioid-involved encounter in NHCS (1). The Co-occurring Disorders Algorithm was designed to complement an existing Enhanced Opioid Identification Algorithm, which was developed for a previous PCORTF project to identify opioid-involvement in a linked hospital and mortality data set (2).

Both algorithms use data science techniques such as regular expression operations and natural language processing (NLP) to determine the occurrence of an event (for example, the use of opioids, type of opioid agent taken, and presence of co-occurring disorders) by searching all available structured and unstructured data. The Co-occurring Disorders Algorithm uses medical codes and additional information from clinical notes, which became available in NHCS for the first time in 2016, to identify and count patients with an opioid-involved encounter with co-occurring disorders. Additionally, the NHCS data are linked to the National Death Index through the NCHS Data Linkage Program, allowing for the assessment of longitudinal outcomes such as postacute mortality. Output from the Co-occurring Disorders Algorithm and the Co-occurring Disorders Dataset is available to researchers through the NCHS and Federal Research Data Centers (RDC). For instructions on submitting a proposal, see the RDC website (https://www.cdc.gov/rdc/index.htm).

This report details the methodology used to create the Co-occurring Disorders Algorithm and the results of applying the algorithm to 2016 NHCS data. Analytic considerations and limitations of the enhanced methodology are also presented, as well as future considerations and uses of the algorithm.

Project Background

The objective of the fiscal year 2019 PCORTF project was to identify the presence of co-occurring disorders among opioid-involved hospital encounters in 2016 NHCS data. Hospitals participating in the 2016 NHCS could submit one of two data sources: 1) Uniform Billing–04 administrative claims data or 2) electronic health records (EHR) data. Submitted data for all hospitals included medical codes (diagnosis and procedure codes, for example) and free text clinical notes for some hospitals that submitted EHR data.

Previous studies have demonstrated that phenotyping algorithms (designed to detect a clinical condition or characteristic) that rely exclusively on searches of medical codes may not be as comprehensive as those supplemented with text mining techniques like NLP applied to clinical notes (3,4). This is particularly true for diagnosis codes, which are primarily used by hospitals for billing purposes rather than research. Additionally, chronic illnesses such as SUDs and MHIs may have been diagnosed before presentation at the hospital, so relevant codes for this pre-existing condition may not appear in medical code lists generated for the current hospital encounter. Searches of the EHR clinical notes sections that include "Past Medical History" and "Social History" can help identify patients who were diagnosed with SUDs or MHIs before the encounter.

Data Source

NHCS is an establishment survey designed to produce national estimates on the characteristics of inpatient hospitalizations and emergency department (ED) encounters, including length of stay, frequency of diagnoses or procedures, demographic characteristics, and patterns of hospital use across the nation (5). The 2016 NHCS sample included 581 noninstitutional nonfederal hospitals in the United States that had six or more staffed inpatient beds. A total of 158 hospitals submitted data in 2016, where 37 hospitals submitted EHR data. The 2016 NHCS includes 7,032,304 total ED and 2,591,722 inpatient encounters. Although the intent of NHCS is to make national estimates, data from the 2016 survey year are not nationally representative due to the low response rate (27.2%). More information on NHCS methodology is published elsewhere (6).

Case Definitions

Case definitions were developed to identify patients experiencing an opioid-involved hospital encounter who also have an SUD, MHI, or both, that is, co-occurring disorders. The case definitions were developed in collaboration with a technical expert panel (TEP). The TEP included representatives from several federal agencies, including NCHS, U.S. Food and Drug Administration, Substance Abuse and Mental Health Services Administration (SAMHSA), National Institute on Drug Abuse, and National Institute of Mental Health. The TEP provided subject-matter expertise on how to define MHI and SUD, classification strategies, medical codes, and search terms that could be used to identify cases.

SUD Encounter

An SUD encounter was defined as an ED visit or hospitalization with a lifetime diagnosis of a use disorder for any of the following substance categories:

- Alcohol
- Cannabis
- Cocaine
- Hallucinogen
- Inhalant
- Opioid
- Other stimulants
- Sedative, hypnotic, or anxiolytic
- Tobacco
- Other psychoactive substance

These substance categories were selected to align with classifications used in Chapter 5 of the *International Classification of Diseases, 10th Revision*, *Clinical Modification* (ICD–10–CM); the "Mental and Behavioral Disorders Due to Psychoactive Substance Use" section (codes F10–F19) of the *International Classification of Diseases, 10th Revision*; and the criteria defined in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*.

Mentions of an SUD could have occurred in the EHR clinical notes or in medical codes associated with the presence of an SUD. Three criteria were established to meet the SUD case definition: 1) presence of at least one selected SUD code in any diagnosis, reason for visit, or problem code field(s); 2) presence of at least one selected SUD in any procedure code field(s); or 3) classification by the NLP processor based on SUD indicators in the text clinical notes.

MHI Encounter

An MHI encounter was defined as an ED visit or hospitalization with a lifetime diagnosis for any of the following selected MHI categories:

- Anxiety—Includes generalized anxiety disorder, panic disorders, social phobias, unspecified anxiety disorder, or other anxiety disorder.
- Depression—Includes major depressive disorder single episode, major depressive disorder recurrent episode, and other depressive disorder. For the purposes of the study, depression comorbidities of bipolar disorder and schizophrenia are excluded.
- Obsessive compulsive disorder—Characterized by persistent and repetitive thoughts that result in repetitive excessive behaviors to reduce repetitive thoughts or mitigate a perceived threat.
- Self-harm—Personal history of self-harm, suicidal ideation, or suicide attempt (7).
- Trauma- and stressor-related disorders—Includes acute stress reaction and post-traumatic stress disorder.

The included MHI categories were not intended to be an exhaustive list. Rather, they were selected in collaboration with the TEP to target specific MHI categories of high priority to each agency and based on previous research that found a higher prevalence of these conditions among patients experiencing an opioid-involved encounter (8,9). Mentions of a selected MHI could have occurred in the EHR clinical notes or in medical codes related to the presence of a selected MHI.

Three criteria were established to meet the MHI case definition: 1) presence of at least one selected MHI code in any diagnosis, reason for visit, or problem code field(s); 2) presence of at least one selected MHI service code in any procedure code field(s); or 3) classification by the NLP processor based on MHI indicators in the text clinical notes.

Co-occurring Disorders Encounter

A co-occurring disorders encounter was defined as an ED visit or hospitalization with evidence of at least one lifetime SUD diagnosis and at least one lifetime MHI diagnosis as defined previously.

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 "Co-occurring Disorders Identification Methodology" section. The final medical code and search term lists can be found in the Co-occurring Disorders Dataset specifications available from: https://www.cdc.gov/nchs/data/nhcs/FY19-RDC-2021-06-01-508.pdf (see Appendixes I–IV).

Co-occurring Disorders Identification Methodology

The Co-occurring Disorders Identification Algorithm consists of two components, the code component and the NLP component, which collectively use all available information collected in NHCS. The code component searches data stored in the medical code fields, while the NLP component uses NLP techniques to search the clinical text note fields. The two components are mutually exclusive and search the appropriate data fields independently.

Code Component Development

Similar to the methodology used for the previous Enhanced Opioid Identification Algorithm (2), a two-phased approach was used to develop the code component for the Co-occurring Disorders Algorithm. First, candidate medical codes and search terms were extracted from existing lists provided by TEP members and other sources for each SUD and MHI category. Next, each list was further refined to match case definition criteria.

Code component phases 1 and 2: Identifying and refining code and search term lists

In collaboration with the TEP, the study team identified existing lists of medical codes and search terms from the following organizations:

- NCHS' Division of Health Care Statistics and SAMHSA—Medical codes for SUDs used in a previous set of algorithms to identify substance-involved ED visits in the 2014 NHCS (10).
- SAMHSA—Key definitions used to identify search terms for the 2016 National Survey on Drug Use and Health (11).
- Agency for Healthcare Research and Quality—Medical codes for depression, drug abuse, and psychoses used in the Elixhauser Comorbidity Software Refined for ICD–10–CM, a tool that identifies pre-existing conditions based on secondary diagnoses (comorbidities) listed in hospital administrative data (12).
- U.S. National Library of Medicine—SUD- and MHI-related service codes and terms from the Value Set Authority Center, a repository of public value sets used to define clinical concepts (13).

During the second phase, relevant medical codes and search terms were extracted and reviewed by study team members trained in pharmacology and emergency medicine and by additional subject-matter experts from the NCHS Clinical Advisory Group. During this process, the initial list of medical codes and search terms was refined to ensure adequate coverage of SUD- and MHI-related diagnoses and procedures to meet all case definition requirements.

Code component: Diagnostic medical codes

Codes drawn from the earlier NCHS substance-involvement algorithms used *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD–9–CM) diagnosis codes to identify use disorders involving 10 priority substance categories (10). In the 2016 NHCS, most diagnostic information was submitted in the newer ICD–10–CM. Consequently, SUD and MHI codes from the original set of algorithms were mapped to the equivalent ICD–10–CM codes. Other ICD–10–CM codes were included as needed to match all SUD and MHI categories described in the "Case Definitions" section.

All diagnostic code fields (diagnosis, reason for visit, and problems) were searched for the final set of ICD–10–CM codes. A small percentage of patients with opioid-involved encounters (11,680 or 0.9% ED and 1,027 or 0.1% inpatient) did not have a diagnosis record with an ICD–10–CM code and were excluded from diagnostic medical code searches. However, these encounters could still be examined for the presence of SUD- or MHI-related procedure codes.

Code component: Procedure codes

Only procedure code fields that explicitly mentioned a definitive SUD or MHI were included in the final algorithm. Healthcare Common Procedure Coding System procedure codes documenting a positive screening test for an SUD or MHI were included. However, codes that were used to indicate that a screening test was conducted, but with no indication of whether the test was positive, were excluded due to the ambiguity of the outcome.

Codes for administration of specific medication-assisted treatment drugs, anxiolytics, and antidepressants were also excluded because these drugs may be prescribed to treat a variety of conditions and symptoms and do not indicate that a patient has a confirmed SUD or MHI of interest. Similarly, codes for behavioral health counseling services were not included because patients may be provided or referred for these services in the absence of a confirmed SUD or MHI diagnosis.

Conducting code component search

The final medical code lists were used to search all available diagnostic and procedure code fields in the 2016 NHCS. SAS 9.4 was used to perform all code-based searches. The final medical code list can be found in the Co-occurring Disorders Dataset specifications available from: https://www.cdc.gov/nchs/data/nhcs/FY19-RDC-2021-06-01-508.pdf (see Appendixes I and III).

NLP Component Development

Algorithms that rely exclusively on medical diagnosis and service codes may miss some hospital encounters involving SUDs and MHIs (10,14). The NLP component was designed to search for other evidence of these conditions in unstructured clinical notes fields. Python 3.7 was used to develop and run NLP processors to examine clinical text.

Annotation of gold standard data set

To develop and test the NLP algorithms, a gold standard data set was created by annotating, or classifying, hospital encounters according to MHI and SUD status. Precisely 1,939 ED and inpatient encounters with at least one clinical note record from the 2016 NHCS were manually reviewed and annotated by a team of clinicians. The development of the annotation data set has been described elsewhere (15). Descriptions of the categories for the fiscal year 2019 project and the number of encounters selected for each are detailed below.

- SUD medical codes—Encounters that have an ICD–10–CM diagnosis code for an SUD. A total of 200 encounters were selected for this category.
- MHI medical codes—Encounters that have an ICD–10–CM diagnosis for an MHI. A total of 100 encounters were selected for this category (50 involving an anxiety disorder and 50 involving a depressive disorder).
- SUD keywords—Encounters that have keyword matches for an SUD but do not have an ICD–10–CM SUD diagnosis code. A total of 300 encounters were selected for this category.
- MHI keywords—Encounters that have keyword matches for an MHI but do not have an ICD–10–CM MHI diagnosis code. A total of 200 encounters were selected for this category.
- Additional encounters—Encounters included for the development of the fiscal year 2018 PCORTF project algorithm identifying any form of opioid use and opioid overdose. A total of 850 encounters were selected for this category.
- A random selection—Implemented to balance the data set (that is, ensure that negative examples were available for training and evaluation) and to include some cases that were relevant but that did not fit into any of the previous encounter categories. The set from which these encounters were selected did not include encounters that had a relevant ICD–10–CM diagnosis code (opioid involved, SUD related, or MHI related) or any encounters that had an opioid term, MHI keyword, or SUD keyword. A total of 300 encounters were selected for this category.

The human-annotated encounters allowed for the comparison of intermediate versions of the algorithms and adjustment of the algorithms to better match clinicians' decisions. The annotated data set was divided into a development set for refinement of the algorithm and a test set for final evaluation of its performance. The outcomes from performance testing are reported in "Results."

Classification of SUD and MHI

The NLP component was designed to search encounters with available clinical notes for mentions of a set of SUD and MHI search terms (either single phrases or phrase combinations) that matched the same concepts in the code component. To define this set, clinicians had to perform upfront exclusions, find SUD and MHI search terms, detect rule-outs, and assign encounters to SUD and MHI categories of interest.

Upfront exclusions based on note type, which represent a different section of a patient's medical chart, were performed to identify and exclude certain out-of-scope note types from the search. These types included immunizations, vital signs, allergies, and patient instructions. The first three note types did not contain information indicating an SUD or MHI diagnosis, while patient instructions provided general information that was not specific to the patient, producing false positives. Searches for SUD and MHI keywords were conducted on all in-scope note types. Note records with an unspecified note type, likely due to a data extraction or submission error, were considered in-scope and included in the search. Additionally, descriptions or labels for diagnosis codes that were embedded in note records were searched if they were not from out-of-scope note types.

Searches were structured to find the strongest evidence of an SUD or MHI first (that is, high-priority terms) and, if not found, to then search for weaker evidence of an SUD or MHI (that is, deprioritized terms). High-priority terms were more specific and less likely to yield false positives. For SUDs, examples of high-priority terms included the standalone phrase "amphetamine use disorder" and the phrase combination of "abuse" with "benzos" in the same sentence (not necessarily adjacent). Deprioritized terms were more general phrases that did not indicate the specific type of SUD, such as "drug abuse," and would only be flagged if no high-priority term was found in the same sentence.

For MHIs, the generic terms "anxiety" and "depression" were considered deprioritized terms compared with more specific phrases like "generalized anxiety disorder," and only flagged if found in note types most likely to contain diagnostic information (for example, chief complaint, problem list, or discharge summary). Terms that refer to temporary or occasional moods like "anxious" or "depressed" were not identified as an MHI. Some keyword matches also were excluded from assignment to the "other depressive disorder" subcategory because they did not meet the case definition criteria. An example is the phrase "ST depression," which is related to heart monitoring rather than a depressive disorder. Similarly, depression-related keyword matches were excluded if they closely occurred with mentions of bipolar disorder or schizophrenia.

Clinical note text was broken up into sentences and searched sentence by sentence to detect several types of rule-outs. SUD and MHI mentions for dates later than 2016 were excluded. Mentions indicating a family history of an SUD or MHI, compared with a personal history, were also excluded. An externally developed negation detection algorithm was also applied, called NegEx, which helps determine whether a finding or disease mentioned within a narrative medical report was present (confirmed) or absent (negated) (16). If an SUD or MHI search term was determined to be negated, it was discarded from the collection of positive results for that encounter. For example, in a sentence like, "pt reports being diagnosed for anxiety but not depression," "anxiety" is a positive match and is captured, while "depression" is negated and consequently discarded.

Once all search terms were flagged and no rule-outs detected, they were mapped to the SUD and MHI categories of interest. The final NLP search term list can be found in the Co-occurring Disorders Dataset specifications available from: https://www.cdc.gov/nchs/data/nhcs/FY19-RDC-2021-06-01-508.pdf (see Appendixes II and IV).

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 Co-occurring Disorders Algorithm. Annotators were asked to identify evidence of SUDs and MHIs, the chart location where this information was found, whether the patient was screened for an SUD or MHI, and whether treatment for an SUD or MHI was initiated during the encounter. The annotators were also instructed to document the exact verbiage in the clinical notes for each SUD and MHI mentioned, noting the subcategory, and to identify if the patient received an SUD- or MHI-related diagnosis code.

The annotation data set was divided into a set to inform the development of the NLP processors (the development set) and a set to evaluate their performance (the evaluation set). The SUD evaluation set included 143 SUD positive and 143 SUD negative encounters as indicated by the annotators. The MHI evaluation set included 50 encounters with no MHI, 50 encounters with an anxiety-related MHI, 50 encounters with a depression-related MHI, and 50 encounters with both an anxiety- and depression-related MHI. In both cases, the development sets included the remainder of the annotation data set (that is, those not included in evaluation sets). The performance of the algorithms was measured against the evaluation sets and is shown in Tables 1–6.

Performance based on results obtained from the code component of the Co-occurring Disorders Algorithm alone is shown in Tables 1 and 2. Performance based on results obtained from the NLP component is shown in Tables 3 and 4. Performance based on results obtained by the full algorithm, including both the code and NLP components, is shown in Tables 5 and 6. Only top-level yes or no questions for identifying evidence of SUDs and MHIs are shown; the results do not include questions regarding chart locations where information was found, clinician assessment for an SUD or MHI, or initiation of treatment for an SUD or MHI during the encounter.

The performance of the various components was assessed using several metrics. Recall (also known as sensitivity or the proportion of true positives correctly classified) is the percentage of true positives over the sum of true positives and false negatives. Precision (also known as positive predictive value or the proportion of classified cases that are true positives) is the percentage of true positives over the sum of true positives and false positives. F1, a common single score for performance, is the harmonic mean of recall and precision. It is meant to provide a balanced measure so that neither recall (sensitivity) nor precision (positive predictive value) is given weight over the other. However, it does not equally account for the true negative rate (specificity) and, so, Matthews correlation coefficient (MCC), identical to Pearson's phi coefficient, is also shown. MCC provides a single score for overall performance in which both positive and negative cases are given equal importance (17).

All calculations are based on numbers found in the confusion matrix, which is a table comparing the gold standard annotator data with the model results used to describe the performance of the classification model, 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.
- The cell for annotator negative and algorithm component negative equals true negatives.

Results

Annotation Results

Results of the annotation reflect the performance of the algorithm in distinguishing between the presence or absence of an SUD or MHI identified by the annotators (Tables 1–6). Results show what would be found if only the code component was used, if only the NLP component was used, or if both components were combined. It is important to note that true cases determined by the annotators include information from both medical codes and clinical notes. Consequently, for either individual component, false negatives may include instances where the relevant information was not available for that component to find; that is, a code component false negative could have had no relevant code but did have relevant information contained in the notes. Conversely, an NLP component false negative could have had no relevant information in the notes but did have relevant information contained in the medical codes.

Performance of the code component

For the code component, there were 130 true positive encounters and 49 true negative encounters for MHIs, with a single false positive and 20 false negatives (Table 1). For SUDs, 131 true positives and 141 true negatives were identified, with 2 false positives and 12 false negatives. Precision was high for identifying MHIs (99.2%) and SUDs (98.5%), indicating a high degree of success in identifying true cases (Table 2). However, the recall was lower for MHIs (86.7%) compared with SUDs (91.6%), indicating that the code component missed some true MHI cases. Both SUDs and MHIs identified by the code component had an F1 greater than 90% (92.5% for MHIs and 94.9% for SUDs). MCC was moderate for MHIs (0.77) but high for SUDs (0.90).

Performance of the NLP component

For the NLP component, there were 112 true positive encounters and 46 true negative encounters for MHIs, with 4 false positives and 38 false negatives (Table 3). For SUDs, 129 true positives and 113 true negatives were identified, with 30 false positives and 14 false negatives. Precision was higher for MHIs (96.6%) than SUDs (81.1%), indicating fewer false positives among MHIs (Table 4). However, recall for SUDs was higher (90.2%) than for MHIs (74.7%), showing that the NLP component of the algorithm missed one out of four true cases of MHIs. F1 scores were similar for MHIs (84.2%) and SUDs (85.4%), but MCC was only moderate for SUDs (0.70) and low for MHIs (0.58) due to the greater proportion of MHI false negatives.

Performance of the Co-occurring Disorders Algorithm

The Co-occurring Disorders Algorithm, which combines both the code and NLP components, had 140 true positive encounters and 46 true negative encounters for MHIs, with 4 false positives and 10 false negatives (Table 5). For SUDs, there were 142 true positive encounters and 113 true negative encounters, with 30 false positives and 1 false negative. Comparing MHIs and SUDs, the precision of the combined algorithm in identifying true cases was higher for MHIs (97.2%) than SUDs (82.6%) (Table 6). For both concepts of interest, recall was high with MHIs at 93.3% and SUDs at 99.3%, and F1 was also high (95.2% for MHIs and 90.2% for SUDs). Overall performance, as measured by MCC, was 0.82 for MHIs and 0.80 for SUDs.

Note that during annotation, long-term opioid use (ICD–10–CM code Z79.891) was included under the SUD opioid use disorder subcategory. However, this concept, including the code and related search terms, was later excluded from the final algorithm because it did not meet the case definition criteria based on the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* classification.

Results of Co-occurring Disorders Algorithm in the 2016 NHCS

In the next stage of analysis, the final Co-occurring Disorders Algorithm was applied to the 2016 NHCS. Within the ED, 28.2% were SUD only, 5.4% were MHI only, and 9.2% were co-occurring disorders (Table 7). Within inpatient hospitalizations, 30.6% were SUD only, 10.0% were MHI only, and 15.0% were co-occurring disorders. Of the 1,370,827 opioid-involved hospital encounters identified, an SUD or MHI was not detected in 57.2% of ED encounters and 44.4% of hospitalizations.

Table 8 shows the number and percentage of ED and inpatient opioid-involved encounters with evidence of an MHI only, an SUD only, or co-occurring disorders by the code component alone, the NLP component alone, and the overlap between the two methodologies (that is, found by both components). Most encounters (75.0%) were flagged by only the code component, while 10.3% were flagged by the NLP component alone, and 14.7% were flagged by both components. This finding reflects, at least in part, the disparate availability of codes (present on 99.4% of encounters) relative to clinical notes (present on 8.7% of encounters).

Table 9 shows the number and percentage of opioid-involved encounters with evidence of an MHI only, an SUD only, or co-occurring disorders identified by each algorithm component by hospital setting. Alone, the code component identified most encounters in both settings but detected a greater percentage of records in the inpatient setting (83.9%) than in the ED setting (66.9%). The NLP component identified a greater percentage of encounters in the ED setting (15.0%) than in the inpatient setting (5.1%). The remaining encounters were identified by both components, including 18.1% in the ED setting and 11.0% in the inpatient setting.

Analytic Considerations and Limitations

The Co-occurring Disorders Dataset was created using 2016 NHCS data, which is not nationally representative because of NHCS' low response rate. Additionally, each of the data sources for the 2016 NHCS (EHR and Uniform Billing–04 administrative claims) have limitations that affected how each component identified encounters with evidence of an SUD and MHI. The following sections describe these limitations, which must be considered to properly interpret the results of applying the final Co-occurring Disorders Algorithm.

Limitations of the code component

While coded information of clinical data generally provides a standardized and efficient way to search for SUDs and MHIs, several hospitals submitted data with incomplete or nonstandardized diagnosis and procedure information. In addition, the number of submitted diagnostic codes differed by data source, with Uniform Billing–04 administrative claims files limited to a maximum of 28 diagnostic codes per encounter (up to 3 in the reason for visit fields and up to 25 in the diagnosis fields). EHR data files, in contrast, could include an unlimited number of diagnostic codes per encounter. Lastly, some diagnoses were embedded within the clinical notes and were only searched by the code component if extracted from the text. It is possible that some embedded diagnoses were missed by the extraction process.

Limitations of the NLP component

The most significant limitation of the NLP component was that only 9.0% of ED encounters and 7.4% of inpatient encounters had any clinical note record available. As a result, the NLP component could not be run on most of the encounters. Among available note records, extraction errors presented challenges for the NLP processors when analyzing the text. For example, all submitted notes from one hospital were shortened to a maximum of 256 characters during extraction. As noted earlier, the detection of rule-outs relies on the ability to first break up the text into sentences. Many notes also had punctuation or white space removed during extraction, which made it difficult to distinguish between sentences.

Lastly, the submitted clinical notes may have excluded some pertinent information stored in psychotherapy notes. The Health Insurance Portability and Accountability Act Privacy Rule states that notes recorded by a mental health professional documenting or analyzing the contents of a conversation during a private counseling session or a group, joint, or family counseling session must be stored separately from the rest of the patient's medical record (18). Psychotherapy notes receive special protections beyond other protected health information, and disclosures of this information are authorized only in limited circumstances. Additionally, participating hospitals may have excluded clinical notes from their behavioral health units based on additional confidentiality protections imposed by state or federal laws, such as Title 42 of the Code of Federal Regulations Part 2 regarding disclosure of information that would identify a person as having or having had an SUD (19).

Discussion

The findings from this report demonstrate that similar methodology used to build the earlier Enhanced Opioid Identification Algorithm can be adapted to identify an important subpopulation of patients experiencing an opioid-involved hospital encounter using the Co-occurring Disorders Algorithm. Both algorithms use all available structured and unstructured data submitted from sampled NHCS hospitals to flag encounters that meet case definitions.

In the 2016 NHCS, 11.6% of all opioid-involved encounters had documentation of co-occurring disorders. When examined by hospital setting, a greater percentage of patients with co-occurring disorders were found in the inpatient setting (15.0%) compared with the ED (9.2%).

When the Co-occurring Disorders Algorithm was compared against the gold standard data set annotated by clinicians, the combined algorithm (both the code and NLP components) performed the best in identifying MHIs by the MCC metric. For SUDs, in contrast, the code component achieved the highest MCC and performed the best overall. Upon further examination, this was discovered to be due largely to the greater proportion of false positive cases resulting from the NLP processors designed for tobacco use disorder, which had a weaker ability to consistently identify the concept of tobacco dependence, and inadequate negation detection to correctly exclude mentions for the denial or absence of tobacco use. In the subsequent application of the combined algorithm to all opioid-involved hospital encounters in the 2016 NHCS, the NLP component was able to detect 10.3% of encounters with evidence of an MHI or SUD that were not identified by the code component.

The combined algorithm had a lower MCC than the code component alone, and this was primarily the result of the NLP component's poorer performance on SUD, particularly with respect to precision (false positives). Although a full error analysis will be conducted in a forthcoming validation study, initial investigation reveals that many of the false positives come from incorrectly flagging a tobacco use disorder. Tobacco use is typically recorded within a "Social History" or equivalent section of an EHR, usually with either structured fields (describing tobacco use type, status, frequency of use, or amount used) or a free text field. The NLP algorithm attempted to extract information about use disorders in the same way for all drugs. It may not have sufficiently accounted for how tobacco use was documented across note records.

Free text regarding tobacco use usually does not contain diagnostic phrases or more descriptive information that more clearly indicates a tobacco use disorder. As a result, a phrase such as "current smoker" would be insufficient standalone evidence and would need to be evaluated with other indicators of dependence such as severity of use. In addition, terms indicating substance used, such as "tobacco" or "smoking," may not have been appropriately evaluated with corresponding negation terms, such as "none" or "never," making this a failure of negation detection resulting from inaccurate grouping of text that was evaluated as a single sentence.

These findings demonstrate the potential of using coded medical data and unstructured clinical notes to identify patients with opioid-involved encounters and co-occurring disorders in NHCS. Without the use of NLP to analyze the EHR clinical notes, 51,594 ED and 15,990 inpatient encounters identified as having a co-occurring disorder would not have been identified. The Co-occurring Disorders Algorithm will be refined and improved in future iterations to better identify opioid-involved hospital encounters with co-occurring disorders.

Although the annotation process was critical in developing the coded and NLP components of the algorithm and measuring initial performance, the clinicians were limited to reviewing information sent by participating hospitals. They did not have access to the full medical charts for each encounter, which may have included additional relevant information that was not extracted for submission. Additionally, the annotation process excluded encounters with no available clinical notes. To provide a more comprehensive measurement of performance, NCHS will be conducting a study to validate both the Enhanced Opioid Identification Algorithm and the Co-occurring Disorders Algorithm. This validation study will involve direct abstraction of full medical charts stored in hospital EHR systems for a sample of encounters flagged by both algorithms, as well as encounters that were not identified as involving opioids, SUDs, or MHIs. The findings of this study will be used to target areas for further refinement to improve the performance of both algorithms. The code used in the refined algorithm is available from GitHub for researchers to apply for similar hospital data.

References

1. U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. Identifying co-occurring disorders among opioid users using linked hospital care and mortality data: Capstone to an existing FY18 OS-PCORTF Project. 2021. Available from: https://aspe.hhs.gov/identifying-co-occurring-disorders-among-opioid-users-using-linked-hospital-care-and-mortality-data-capstone-existing-fy18-os-pcortf-project.
2. U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. Enhancing identification of opioid-involved health outcomes using linked hospital care and mortality data. 2021. Available from: https://aspe.hhs.gov/enhancing-identification-opioid-involved-health-outcomes-using-linked-hospital-care-and-mortality-data.
3. Wei WQ, Teixeira PL, Mo H, Cronin RM, Warner JL, Denny JC. Combining billing codes, clinical notes, and medications from electronic health records provides superior phenotyping performance. J Am Med Inform Assoc 23(e1):e20–7. 2016.
4. Smoller JW. The use of electronic health records for psychiatric phenotyping and genomics. Am J Med Genet B Neuropsychiatr Genet 177(7):601–12. 2018.
5. National Center for Health Statistics. National Hospital Care Survey. 2021. Available from: https://www.cdc.gov/nchs/nhcs/index.htm.
6. National Center for Health Statistics. The linkage of the 2016 National Hospital Care Survey to the 2016/2017 National Death Index: Methodology overview and analytic considerations. 2019. Available from: https://

www.cdc.gov/nchs/data/datalinkage/NHCS16_NDI16_17_Methodology_Analytic_Consider.pdf.

7. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. 2013. Available from: https://acrobat.adobe.com/link/track?uri=urn%3Aaaid%3Ascds%3AUS%3A907fa51f-b6cb-494c-95b1-5cacf626fc55#pageNum=1.

8. Khazaal Y, Krenz S, Benmebarek M, Zullino DF. Worsening of obsessive-compulsive symptoms under methadone tapering. Prog Neuropsychopharmacol Biol Psychiatry 30(7):1350–2. 2006.

9. Davis MA, Lin LA, Liu H, Sites BD. Prescription opioid use among adults with mental health disorders in the United States. J Am Board Fam Med 30(4):407–17. 2017.

10. 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.

11. Substance Abuse and Mental Health Services Administration. 2016 National Survey on Drug Use and Health: Methodological summary and definitions. 2017.

12. U.S. Department of Health and Human Services. Elixhauser comorbidity software refined for ICD–10–CM. 2020. Available from: https://www.hcup-us.ahrq.gov/toolssoftware/comorbidityicd10/comorbidity_icd10.jsp.

13. National Institutes of Health. Value set authority center. 2021. Available from: https://vsac.nlm.nih.gov/.

14. Ridgway JP, Uvin A, Schmitt J, Oliwa T, Almirol E, Devlin S, Schneider J. Natural language processing of clinical notes to identify mental illness and substance use among people living with HIV: Retrospective cohort study. JMIR Med Inform 9(3):e23456. 2021.

15. 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.

16. 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. 2001.

17. Chicco D, Tötsch N, Jurman G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Mining 14(13). 2021. Available from: https://doi.org/10.1186/s13040-021-00244-z.

18. U.S. Department of Health and Human Services. HIPAA privacy rule and sharing information related to mental health. 2021. Available from: https://www.hhs.gov/sites/default/files/hipaa-privacy-rule-and-sharing-info-related-to-mental-health.pdf.

19. Substance Abuse and Mental Health Services Administration. Substance abuse confidentiality regulations. 2021. Available from: https://www.samhsa.gov/about-us/who-we-are/laws-regulations/confidentiality-regulations-faqs.

Table 1. Agreement counts between the code component of the Co-occurring Disorders Algorithm for mental health issues and substance use disorders and the annotated data set

Algorithm	Mental health issue encounter: Annotator positive	Mental health issue encounter: Annotator negative	Substance use disorder encounter: Annotator positive	Substance use disorder encounter: Annotator negative
Algorithm positive	130	1	131	2
Algorithm negative	20	49	12	141

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

Table 2. Performance measures of the code component of the Co-occurring Disorders Algorithm and the annotated data set

Measure	Mental health issue	Substance use disorder
Recall[1]	86.7	91.6
Precision[2]	99.2	98.5
F1[3]	92.5	94.9
Matthews correlation coefficient	0.77	0.90

[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 3. Agreement counts between the natural language processing component of the Co-occurring Disorders Algorithm for mental health issues and substance use disorders and the annotated data set

	Mental health issue encounter		Substance use disorder encounter	
Algorithm	Annotator positive	Annotator negative	Annotator positive	Annotator negative
Algorithm positive	112	4	129	30
Algorithm negative	38	46	14	113

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

Table 4. Performance measures of the natural language processing component of the Co-occurring Disorders Algorithm and the annotated data set

Measure	Mental health issue	Substance use disorder
Recall[1]	74.7	90.2
Precision[2]	96.6	81.1
F1[3]	84.2	85.4
Matthews correlation coefficient	0.58	0.70

[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 5. Agreement counts between the Co-occurring Disorders Algorithm for mental health issues and substance use disorders and the annotated data set

Algorithm	Mental health issue encounter		Substance use disorder encounter	
	Annotator positive	Annotator negative	Annotator positive	Annotator negative
Algorithm positive	140	4	142	30
Algorithm negative	10	46	1	113

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

Table 6. Performance measures of the Co-occurring Disorders Algorithm and the annotated data set

Measure	Mental health issue	Substance use disorder
Recall[1]	93.3	99.3
Precision[2]	97.2	82.6
F1[3]	95.2	90.2
Matthews correlation coefficient	0.82	0.80

[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. Number and percentage of opioid-involved hospital encounters, by presence of substance use disorder and mental health issue

Presence of SUD and MHI	Emergency department encounter		Inpatient encounter	
	Number	Percent	Number	Percent
No SUD or MHI	460,677	57.2	250,925	44.4
SUD only	226,746	28.2	172,886	30.6
MHI only	43,561	5.4	56,541	10.0
Co-occurring disorders	74,472	9.2	85,019	15.0
Total encounters	805,456	100.0	565,371	100.0

NOTES: SUD is substance use disorder. MHI is mental health issue. Data are unweighted and are not nationally representative.

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

Table 8. Number and percentage of opioid-involved hospital encounters with MHIs only, SUDs only, and co-occurring disorders identified by the code, NLP, and both components of the Co-occurring Disorders Algorithm

Algorithm component	Number	Percent
Code component only	494,458	75.0
NLP component only	67,584	10.3
Both components	97,183	14.7
Total encounters	659,225	100.0

NOTES: MHI is mental health issue. SUD is substance use disorder. NLP is natural language processing. Data are unweighted and are not nationally representative.

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

Table 9. Number and percentage of opioid-involved hospital encounters with MHIs only, SUDs only, and co-occurring disorders identified by the code, NLP, and both components of the Co-occurring Disorders Algorithm, by setting

	Emergency department encounter		Inpatient encounter	
Algorithm component	Number	Percent	Number	Percent
Code component only	230,612	66.9	263,846	83.9
NLP component only	51,594	15.0	15,990	5.1
Both components	62,573	18.1	34,610	11.0
Total encounters	344,779	100.0	314,446	100.0

NOTES: MHI is mental health issue. SUD is substance use disorder. NLP is natural language processing. Data are unweighted and are not nationally representative.

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. 193
CS332608