U.S. Department of Justice
Office of Justice Programs
Bureau of Justice Statistics

TECHNICAL REPORT

April 2016, NCJ 249568

Assessing Inmate Cause of Death: Deaths in Custody Reporting Program and National Death Index

Zhen Zeng, Ph.D., Margaret Noonan, and E. Ann Carson, Ph.D., *BJS Statisticians*
Ingrid Binswanger, M.D., *BJS Visiting Fellow*
Patrick Blatchford, Ph.D., *Colorado School of Public Health*
Hope Smiley-McDonald, Ph.D., and Chris Ellis, *RTI International*

Introduction

The U.S. Department of Justice's Bureau of Justice Statistics (BJS) has collected data annually on inmates who died in state prison and local jail and the circumstances surrounding these deaths since the Death in Custody Reporting Act (P. L. 106–297) was passed in 2000. There were no standardized requirements for prison and jail administrators to report inmate deaths prior to the passage of the act.

Under BJS's Deaths in Custody Reporting Program (DCRP), respondents from jails and state departments of corrections are asked to report causes of deaths (COD), identifiers, and characteristics of each inmate who died. This includes personal data (i.e., name, date of birth, date of death, and sex), correctional information (i.e., length of stay, legal status, and criminal offense of the deceased), and the circumstances of the death (i.e., medical treatment before death, pre-existing conditions, and time and place of death incident).

The DCRP had a 100% response rate among 50 state departments of corrections and 99% among the approximately 2,800 jail jurisdictions. Item response rates were also generally close to 100%. However, BJS had some concerns about the collection of COD data in the DCRP. Due to changes in data collection agents in 2008, data quality follow-up was cut short, resulting in 21% of missing COD information for jail deaths that year, compared to 2% in 2007 and 2009. This data anomaly significantly compromised BJS's ability to report death trends.

Although the DCRP has collected inmate COD data since 2000, the extent to which respondents accurately reported this information, especially in multiple-cause deaths, is unknown. Various correctional professionals may complete the DCRP records, including jail administrators and correctional medical and line staff, while only the treating physician, a forensic pathologist, coroner, or funeral director may complete and file death certificates in the United States.

The DCRP did not collect information on the manner of death for suicides (e.g., hanging) until 2008 and for alcohol and drug intoxication (e.g., ethanol toxicity and heroin toxicity) until 2009.

DCRP respondents are instructed to report the final COD (known as the immediate COD on the death certificate) as recorded by a medical examiner or through another official medical investigation. (See *Collecting cause of death data in the DCRP and NDI* textbox.) In comparison, U.S. death certificates capture "the chain of events—diseases, injuries, or complications—that directly caused the death" ranked in order from the immediate to underlying. Contributing causes, which did not result in the underlying cause, are also collected in the National Center for Health Statistics' (NCHS) National Death Index (NDI).

COD is one of the items that require the highest level of data quality follow-up in the DCRP collection due to item nonresponse or invalid data. If an autopsy is requested—as it is for nearly all jail and most prison deaths—DCRP respondents must wait for the autopsy report and all associated toxicology and lab tests to be made available. In some cases, respondents cannot obtain COD information from the medical examiner and will submit a record that is missing the COD. Missing data require follow-up and can delay the final death record for weeks or months.

To address these issues, BJS examined the NDI as an alternative means of capturing COD data. The NDI is a centralized database of death certificate information abstracted from state vital statistics offices. Through NDI searches, health and medical researchers can determine whether persons in their studies have died or establish COD for persons known to be deceased. The NDI is considered the most comprehensive and accurate collection of death identification and COD information. Both the source (e.g., death certificates) and coverage are better than other mortality death files, such as the Social Security Administration's Death Master File.[1]

[1]Fillenbaum, G. G., Burchett, B. M., & Blazer, D. G. (2009). Identifying a National Death Index match. *American Journal of Epidemiology, 170*(4), 515-518.

BJS had several objectives for this project: obtain the COD codes reported on death certificates, evaluate the quality of the matches, and assess the agreement in COD between the matched DCRP and NDI records.

This report summarizes the results of these analyses and provides recommendations for improving COD data by linking DCRP and NDI data.

DCRP and NDI cause of death coding

Both the DCRP and NDI use codes for COD from the 10th revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10). However, the two collections differ in how they assign ICD-10 codes and identify the most important COD when multiple conditions are present. These differences play an important role in COD mismatch between the DCRP and NDI **(table 1)**.

A nosologist (i.e., a clinical coder trained by the NCHS) processes DCRP data based on the codes listed on DCRP death forms. The nosologist assigns a maximum of five codes to each DCRP death case. The NDI provides up to 20 ICD-10 codes for the various conditions a decedent had, with a maximum of 14 for this dataset. While the NDI assigns ICD-10 codes for all deaths, the DCRP assigns ICD-10 codes only to natural deaths, such as heart disease, cancer, and liver disease. DCRP respondents are instructed to check a box for AIDS-related deaths, but no additional COD information is captured for these deaths. Unnatural deaths, such as drug or alcohol intoxication, homicides, suicides, and accidents, are not assigned ICD codes in the DCRP.

The NDI uses the Automated Classification of Medical Entities (ACME) program to identify the underlying COD, which is the international standard for automated selection of the underlying COD. Studies have shown little discrepancy (up to 7%) between underlying causes assigned by ACME and a professional nosologist, but ACME offers considerable savings both in cost and efficiency.[2]

The inputs to ACME are the multiple ICD codes assigned to each entity (e.g., disease condition, accident, or injury) listed on death certificates, in a particular order as reported. ACME then applies the World Health Organization rules to the ICD codes and selects an underlying COD. Underlying COD is defined as "(a) the disease or injury which initiated the train of events leading directly to death, or (b) the circumstances of the accident or violence which produced the fatal injury."[3]

The DCRP uses a different system for identifying the underlying COD. Specifically, respondents are instructed to report a final COD as determined by a medical examiner. Cases with a single COD are ranked by frequency separately for male and female decedents. When a death is caused by multiple illnesses, the most common sex-specific COD among those with only one COD is designated the underlying COD. For example, a male decedent with two conditions—liver disease and cancer—is counted as a cancer death because cancer is a more common COD than liver disease among men who die from a single COD.

[2] Doody, M. M., Hayes, H. M., & Bilgrad, R. (2001). Comparability of National Death Index Plus and standard procedures for determining causes of death in epidemiologic studies. *Annals of Epidemiology, 11*(1), 46-50.

[3] World Health Organization. (1975). *Manual of the international statistical classification of diseases, injuries, and causes of death: World Health Organization*.

TABLE 1
Comparison of DCRP and NDI cause of death data

	DCRP	NDI
Data source	DCRP death forms	Death certificates
Data provider	Prison and jail administrator or correctional medical staff	Treating physician, forensic pathologist, coroner, or funeral director
Availability of ICD-10 codes	Illness deaths only	All deaths
Maximum number of ICD-10 codes	5	14[a]
Selection of the underlying COD	Most common cause for single-cause deaths[b]	Automated Classification of Medical Entities (ACME) program
Categorization of the underlying COD	11 major categories of death	ICD-10 code, 113 categories, and 11 major categories of death
Time lag	9 months from end of calendar year	12 months from end of calendar year

Note: COD = cause of death.

[a]The NDI records a maximum of 20 International Classification of Diseases (ICD-10) codes. The matched DCRP 2007–10 death records captured a maximum of 14 codes.

[b]Cases with a single COD are ranked by frequency separately for male and female decedents. When a death is caused by multiple illnesses, the most common sex-specific COD among those with only one COD is designated the underlying COD.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

In the NDI, the underlying COD variable is assigned an ICD-10 code. This variable is collapsed into broader COD categories: first into a 113-category Level-2 COD variable and then into an 11-category Level-1 COD variable. For instance, ICD-10 codes E10 through E14 (i.e., insulin-dependent diabetes mellitus, non-insulin-dependent diabetes mellitus, malnutrition-related diabetes mellitus, other specified diabetes mellitus, and unspecified diabetes mellitus) are grouped into diabetes mellitus, a Level-2 category. The Level-1 CODs include 11 major categories of death:

- heart disease
- AIDS-related
- cancer
- liver disease
- respiratory disease
- other illness
- suicide
- homicide
- drug and alcohol intoxication
- accident
- other.

Both Level-1 and Level-2 categories are standard classifications of CODs that are used by epidemiologists to categorize CODs and by the Centers for Disease Control and Prevention (CDC) and the NCHS.

In the DCRP, the underlying COD is a categorical variable with the same 11 groups as the Level-1 underlying COD in the NDI; the DCRP does not offer ICD-10 codes or detailed categories for the underlying COD. Because BJS publishes inmate death rates by major categories of COD annually, this study focuses on the 11-category underlying COD variable when comparing DCRP and NDI data. Comparisons of multiple causes using all available ICD-10 codes were also conducted.

Collecting cause of death data in the DCRP and NDI

DCRP form

15. What was the cause of death? * Please SPECIFY cause of death—it is critical information *****

- Illness—Exclude AIDS-related deaths [*Specify*] _____
- Acquired Immune Deficiency Syndrome (AIDS)
- Accidental alcohol/drug intoxication [*Describe*] _____
- Accidental injury to self [*Describe*] _____
- Accidental injury by other (e.g., vehicular accidents during transport) [*Describe*] _____
- Suicide (e.g., hanging, knife/cutting instrument, intentional drug overdose) [*Describe*] _____
- Homicide [*Describe*] _____
- Other cause(s) [*Specify*] _____

U.S. standard certificate of death

CAUSE OF DEATH (See instructions and examples)

32. **PART I.** Enter the chain of events—diseases, injuries, or complications—that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.

IMMEDIATE CAUSE (Final disease or condition resulting in death ---------> a. _____

Sequentially list conditions, if any, leading to the cause listed on line a. Enter the **UNDERLYING CAUSE** (disease or injury that initiated the events resulting in death) **LAST.**

b. _____

c. _____

d. _____

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I.

Matching procedure

DCRP death records from 2007 through 2010 were sent to the NDI Plus service, which provides COD listings for each potential match. Subjects in the DCRP were considered known decedents by the NDI, meaning that the DCRP had established the fact of death. Known decedent searches result in a higher likelihood of true matches than searches for subjects with an unknown vital status.

The NDI prefers using Social Security numbers (SSNs) as identifiers. Neither SSNs nor maiden names of female inmates are collected by DCRP and could be linked with the NDI. The match between DCRP and NDI records used the following variables: decedents' birth and death dates, first and last names, sex, race or Hispanic origin, and state of residence.

The DCRP 2007–10 file contained 17,420 records. Analysis was limited to those records with complete information on birth and death dates and unique birth date and death date combinations. Some pairs of subjects with the same birth and death dates also matched on last name, sex, race or Hispanic origin, and state of residence, differing only in first name. It is likely that these were duplicate records, and the difference in the first names was due to the inmate's use of an alias. The final study cohort contained 17,358 decedents (figure 1).

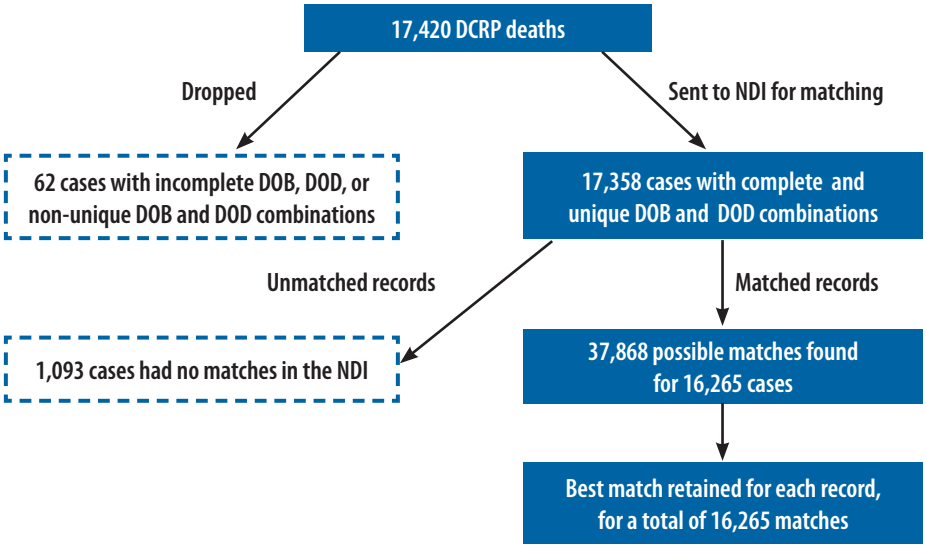
The search for 17,358 DCRP subjects in the NDI database resulted in 37,868 possible matches. Each possible match was assigned a probabilistic match score by the NDI system. The probabilistic score is calculated by summing the individual weights applied to each data element (e.g., decedents' sex, race or Hispanic origin, date of birth, state of residence) used in the NDI match. The weights reflect the degree of agreement between the information on the submission record and the NDI record.[4] For subjects with multiple possible matches, the record with the highest probabilistic score was selected. This procedure reduced the total number of matches to 16,265 deaths, or one match per subject.

In addition to the probabilistic match score, NDI provides two variables to assess the quality of potential matches: status code and exact match indicator. Derived from the probabilistic match score, the status code indicates whether an individual is either assumed to be alive (status code=0) or dead (status code=1). In this study, given that all DCRP subjects are known to be dead, a status code of 0 means that the potential match has a relatively low probability of being a true match. The NDI assigns status codes conservatively to avoid incorrectly assuming that a living person is dead. As a result, while most of the potential matches assigned a status code of 1 were true matches, a number of true matches were assigned a status code of 0.

The exact match variable indicates that all match fields on the user record agree exactly with the information in the NDI. Exact match is a more conservative criterion than status code and may misclassify more true matches (i.e., false negatives) due to misspellings or nicknames (e.g., Robert compared to Bob). Given that all subjects were known to be dead, a broad definition of match was employed to avoid the risk of dropping eligible deaths, and all 16,265 cases with NDI matches were included in the analysis.

[4] National Center for Health Statistics. (2013). *National Death Index user's guide*. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

FIGURE 1
DCRP–NDI data linkage flowchart



Note: DOB = date of birth; DOD = date of death.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

Quality of matching

Overall, the NDI had at least one record available for 94% of the DCRP subjects. No matches were found for 1,093 DCRP records (table 2). Eighty-two percent of the records had a status code of 1 and 67% had an exact match.

The overall match rate (94%) was slightly lower than in other samples, such as Vietnam era veterans (97%) and refinery and petrochemical workers (97%).[5] However, in the absence of several important identifiers (e.g., SSN and maiden name for females), this was a decent match rate.

[5] Boyle, C. A. & Decoufle, P. (1990). National sources of vital status information: Extent of coverage and possible selectivity in reporting. *American Journal of Epidemiology, 131*(1), 160-168; and Acquavella, J. F., Donaleski, D., & Hanis, N. M. (1986). An analysis of mortality follow-up through the National Death Index for a cohort of refinery and petrochemical workers. *American Journal of Indian Medicine, 9*(2), 181-187.

TABLE 2
Matching results between the DCRP and NDI, 2007–2010

	N	NDI matched	Status code = 1[a]	Exact match[b]
Total	17,358	93.7%	81.9%	66.6%
Facility type				
State prison	13,470	94.2%	82.2%	67.8%
Local jail	3,888	91.9	81.0	62.4
Year				
2007	4,483	92.7%	79.5%	69.1%
2008	4,397	93.2	81.5	64.9
2009	4,359	94.7	83.2	63.8
2010	4,119	94.2	83.5	68.5
Sex				
Male	16,321	93.9%	81.9%	66.9%
Female	1,037	90.3	82.2	61.7
Race/Hispanic origin				
White[c]	9,089	94.3%	83.1%	71.2%
Black/African American[c]	5,930	93.6	81.3	63.7
Hispanic/Latino	1,940	91.6	78.4	60.6
Other[c,d]	354	92.1	81.6	29.1
Age				
17 or younger	20	90.0%	80.0%	55.0%
18–24	625	91.8	85.9	64.0
25–34	1,599	92.5	87.4	65.1
35–44	2,803	92.9	82.5	66.3
45–54	5,179	93.6	80.9	66.2
55 or older	7,132	94.5	80.9	67.5

Note: Detail may not sum to total due to missing cases.

[a] The NDI assigns a status code for each matched record. A code of 1 indicates a high probabilistic match score, and a code of 0 indicates otherwise.

[b] Exact match means that all items provided on the user record agree exactly with the items in the NDI.

[c] Excludes persons of Hispanic or Latino origin.

[d] Includes American Indians or Alaska Natives; Asians, Native Hawaiians, or Other Pacific Islanders; and persons of two or more races.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

The match rate did not show much variation by facility type, decedent's sex, race or Hispanic origin, age, or death year. Even when differences were statistically significant between groups (e.g., prison compared to jail deaths and male compared to female decedents), they amounted to about 2 to 4 percentage points.

Comparing data coverage for cause of death

About 2% of death records in 2007–10 were missing COD in the DCRP, compared to 6% in the NDI (table 3). COD data may be missing in the NDI for a DCRP record for two reasons: (1) no matched record was found in the NDI, or (2) a matched record was found but the NDI had no COD information for it. In this comparison, all missing COD in the NDI was due to unmatched records. All matched records had COD data in the NDI. Out of the 336 cases with missing COD in the DCRP, COD data was available for 302 cases by linking with the NDI, leaving 34 cases with no COD information (not shown).

The rate of missing data was higher for jail deaths than prison deaths, particularly in the DCRP (6% missing in jail compared to 1% in prison). This was due to a data anomaly in 2008, when 21% of the DCRP jail deaths (n=202) were missing COD information.[6] By linking to the NDI database, COD information was obtained for 185 out of those 202 cases, providing a resolution for 92% of the cases (not shown).

[6] Excluding 2008, 1% of jail deaths were missing COD.

TABLE 3
Percent missing cause of death data in the DCRP and NDI, by facility type, 2007–2010

	DCRP	NDI*
Total	1.9%	6.3%
State prison	0.7%	5.8%
2007	1.5	6.8
2008	0.0	6.4
2009	1.0	4.6
2010	0.1	5.3
Local jail	6.3%	8.1%
2007	1.8	8.8
2008	21.2	8.0
2009	2.1	7.8
2010	0.4	7.6

*Missing cause of death in the NDI is due to unmatched records. All matched records in the NDI have cause of death information.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

The NDI provided more ICD-10 codes than the DCRP. On average, each subject received 1.06 ICD codes in the DCRP (including unmatched cases), while each matched record received an average of 2.69 codes in the NDI (table 4). This large difference occurred in part because the DCRP asks respondents for final cause of death, and the cause of death question on the U.S. standard certificate of death requests multiple causes. In addition, ICD-10 codes were not assigned to unnatural deaths in the DCRP.

Prison deaths were assigned more ICD codes than jail deaths both in the DCRP (1.17 for prison deaths compared to 0.68 for jail deaths) and in the NDI (2.74 for prison deaths compared to 2.54 for jail deaths). Large differences also existed by COD category. The DCRP did not assign ICD codes for unnatural deaths, such as suicides, homicides, intoxications, and accidents. In the NDI, deaths from unnatural causes received more ICD codes than deaths from illness.

Comparing underlying cause of death

In this report, DCRP and NDI data are compared to determine the accuracy and completeness of CODs reported by DCRP respondents. COD data were compared in three ways. The distributions of underlying COD in the two datasets were compared at the aggregate level. Also, the correspondence of underlying COD between the two datasets was examined at the individual level, and the reasons for mismatch were explored. The first two analyses used the major category of death, which classified COD into 11 groups. Finally, the analysis compared the correspondence of multiple causes between the two datasets at the individual level, using all available ICD-10 codes. This analysis helped explain the mismatch of the underlying COD in the two datasets.

TABLE 4

Mean number of ICD codes available in the DCRP and NDI, 2007–2010

	DCRP		NDI	
	Mean	Standard deviation	Mean	Standard deviation
Total	1.06	0.95	2.69	1.63
Facility type				
State prison	1.17	0.93	2.74	1.69
Local jail	0.68	0.95	2.54	1.40
Status code[a]				
0	1.08	0.97	2.83	1.65
1	1.06	0.95	2.67	1.63
Sex				
Male	1.07	0.95	2.69	1.63
Female	0.97	0.99	2.78	1.70
Year				
2007	1.00	0.93	2.67	1.60
2008	1.01	0.93	2.68	1.61
2009	1.10	0.96	2.74	1.68
2010	1.15	0.99	2.69	1.66
Cause of death				
Heart disease	1.50	0.90	2.72	1.62
AIDS-related	0.49	1.14	2.18	1.58
Cancer	1.35	0.78	2.34	1.60
Liver disease	1.57	0.93	2.76	1.64
Respiratory disease	1.34	0.72	2.80	1.64
Other illness[b]	1.07	0.84	3.00	1.81
Suicide	:	:	2.34	0.86
Homicide	:	:	3.17	1.47
Drug/alcohol intoxication	:	:	3.65	1.49
Accident	:	:	3.60	1.78
Other/unknown[c]	:	:	3.62	1.69

Note: The mean number of International Classification of Diseases (ICD-10) codes in the DCRP is calculated by the DCRP cause of death category, while the mean number in the NDI is calculated by the NDI cause of death category.

: Not calculated. Unnatural deaths in the DCRP were not assigned ICD-10 codes.

[a]The NDI assigns a status code for each matched record. A status code of 1 indicates a high probabilistic match score, and a code of 0 indicates otherwise.

[b]Includes illnesses such as cerebrovascular disease, diabetes, and other nonleading natural causes of death.

[c]Includes cases where cause of death was provided but the manner was not known, cases where an autopsy was inconclusive, and cases where the information listed was a symptom rather than a cause or manner of death.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

The marginal distributions of the underlying COD from the two datasets were similar. The leading CODs in the DCRP were heart disease (25%), cancer (22%), and suicide (11%), while those in the NDI were cancer (23%), heart disease (22%), and suicide (11%) (table 5). Illness, or natural, deaths were far more prevalent among prisoners (89% of prison deaths in the DCRP and the NDI) than among jail inmates (55% of jail deaths in the DCRP and 59% in the NDI). About 30% of jail deaths were due to suicide, compared to 6% of prison deaths.

The leading CODs were the same for male and female inmates (i.e., heart disease, cancer, and suicide) with small differences: Cancer and heart disease accounted for higher percentages of male deaths than female deaths, while suicide accounted for a higher percentage of female deaths than male deaths.[7] These patterns were evident in both the DCRP and the NDI.

To examine how COD data in the NDI correspond to data in the DCRP at the individual level, the DCRP underlying COD was cross-tabulated by the NDI underlying COD for 15,963 respondents with known COD in both datasets. This reveals how frequently the COD categorization of NDI agreed with that of DCRP (table 6 and table 7).

Two statistics are used to measure the congruence of DCRP and NDI COD data (i.e., intercoder reliability): the agreement rate and the kappa statistic. The agreement rate is the percentage of records classified into the same category by the two datasets. The kappa statistic was first proposed by Cohen.[8] It measures intercoder reliability while taking into account agreement that is expected purely by chance—that is, it indicates the level of agreement achieved beyond chance. When the raters (i.e., the DCRP and NDI) are in complete agreement, kappa takes on the value of 1. Landis and Koch suggested characterizing values less than 0 as indicating no agreement, 0 to 0.20 as slight, 0.21 to 0.40 as fair, 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial, and 0.81 to 1 as almost perfect agreement.[9] (See *Methodology* for more detail.)

The overall agreement rate was 70%, meaning that 70% of the deaths were classified into the same major category of death by the DCRP and the NDI. The kappa statistic was 0.64, which indicates a substantial level of agreement following Landis and Koch's guidelines. The percentage of records that received identical classification in both collections varied by COD category, ranging from less than 40% to more than 90%. The categories in the DCRP with the highest agreement rates were suicide (93%), cancer (87%), homicide (86%), and AIDS-related (79%). In the DCRP, deaths due to liver disease (38%) and respiratory disease (48%) had the lowest agreement rates. This low agreement rate was likely due to how DCRP respondents recorded these deaths. For example, a death due to hepatitis (an inflammation of the liver caused by the hepatitis virus) would be recorded accurately as an infectious disease by the NDI, but may have been recorded as a liver-related death by the DCRP. Similarly, respiratory deaths are frequently coupled with other conditions and diseases (such as AIDS, cancer, or heart disease), and it is likely that DCRP respondents listed the causes that they deemed more serious than others.

[7] Although suicides accounted for a larger percentage of female inmate deaths, males accounted for the majority (91% in jail and 95% in prison) of suicides in correctional facilities. See *Mortality in Local Jails and State Prisons, 2000–2013 - Statistical Tables* (NCJ 248756, BJS web, August 2015).

[8] Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46.

[9] Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

TABLE 5
Distributions of cause of death in the DCRP and NDI, by facility type and sex of decedent, 2007–10

	All		State prison		Local jail		Male		Female	
	DCRP	NDI	DCRP	NDI	DCRP	NDI	DCRP	NDI	DCRP	NDI
All causes	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Illness	81.8%	82.8%	88.9%	89.0%	55.2%	59.3%	82.1%	83.0%	77.9%	79.4%
Heart disease	24.5	22.3	24.9	22.0	23.3	23.4	24.8	22.5	20.3	17.6
AIDS-related	3.0	3.7	2.8	3.9	3.4	3.3	2.9	3.7	4.7	4.5
Cancer	21.8	23.2	26.6	28.0	4.0	5.0	22.3	23.5	15.0	17.1
Liver disease	8.1	5.1	9.3	5.6	3.7	3.3	8.3	5.2	4.7	4.8
Respiratory disease	5.9	5.1	6.5	5.6	3.6	3.5	5.8	5.1	7.4	5.4
Other illness[a]	18.5	23.3	18.8	24.0	17.3	20.8	18.0	22.9	25.9	30.0
Unnatural deaths	16.8%	16.6%	10.1%	10.4%	41.6%	39.6%	16.6%	16.4%	20.2%	19.6%
Suicide	11.4	10.9	6.2	5.9	31.0	29.6	11.2	10.7	14.1	13.1
Homicide	1.8	1.9	1.7	1.8	2.0	1.9	1.9	2.0	0.1	0.1
Drug/alcohol intoxication	2.5	2.3	1.4	1.4	6.5	5.5	2.4	2.1	4.7	4.4
Accident	1.1	1.6	0.9	1.3	2.1	2.7	1.1	1.6	1.3	2.0
Other/unknown[b]	1.4%	0.7%	1.0%	0.6%	3.2%	1.1%	1.4%	0.7%	2.0%	1.0%

Note: Includes only records with a known underlying cause of death in both the DCRP and NDI (n=15,963).

[a] Includes illnesses such as cerebrovascular disease, diabetes, and other nonleading natural causes of death.

[b] Includes cases where cause of death was provided but the manner was not known, cases where an autopsy was inconclusive, and cases where the information listed was a symptom rather than a cause or manner of death.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

TABLE 6
Number of NDI causes of death, by DCRP cause of death category, 2007–2010

DCRP cause of death	Total	Heart disease	AIDS-related	Cancer	Liver disease	Respiratory disease	Other illness[a]	Suicide	Homicide	Drug/alcohol intoxication	Accident	Other/unknown[b]
All causes	15,963	3,553	597	3,696	820	821	3,724	1,732	295	361	256	108
Heart disease	3,917	2,745	27	146	59	148	722	7	9	21	24	9
AIDS-related	471	22	371	15	9	8	41	0	0	1	4	0
Cancer	3,487	101	33	3,046	58	42	177	8	4	6	10	2
Liver disease	1,296	38	34	65	497	12	630	2	0	6	8	4
Respiratory disease	943	99	38	83	35	449	218	0	3	2	11	5
Other illness[a]	2,946	447	87	290	144	130	1,739	8	21	24	40	16
Suicide	1,816	12	3	11	2	10	46	1,679	9	11	25	8
Homicide	279	9	0	6	0	2	9	0	240	0	5	8
Drug/alcohol intoxication	398	26	0	3	2	4	41	12	3	259	4	44
Accident	181	6	0	5	1	1	18	3	5	19	116	7
Other/unknown[b]	229	48	4	26	13	15	83	13	1	12	9	5

Note: Includes only records with a known underlying cause of death in both the DCRP and NDI (n=15,963).

[a]Includes illnesses such as cerebrovascular disease, diabetes, and other nonleading natural causes of death.

[b]Includes cases where cause of death was provided but the manner was not known, cases where an autopsy was inconclusive, and cases where the information listed was a symptom rather than a cause or manner of death.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

TABLE 7
Percent of NDI causes of death, by DCRP cause of death category, 2007–2010

DCRP cause of death	Total	Heart disease	AIDS-related	Cancer	Liver disease	Respiratory disease	Other illness[a]	Suicide	Homicide	Drug/alcohol intoxication	Accident	Other/unknown[b]
All causes	100%	22.3%	3.7%	23.2%	5.1%	5.1%	23.3%	10.9%	1.9%	2.3%	1.6%	0.7%
Heart disease	100%	70.1	0.7	3.7	1.5	3.8	18.4	0.2	0.2	0.5	0.6	0.2
AIDS-related	100%	4.7	78.8	3.2	1.9	1.7	8.7	0.0	0.0	0.2	0.9	0.0
Cancer	100%	2.9	1.0	87.4	1.7	1.2	5.1	0.2	0.1	0.2	0.3	0.1
Liver disease	100%	2.9	2.6	5.0	38.4	0.9	48.6	0.2	0.0	0.5	0.6	0.3
Respiratory disease	100%	10.5	4.0	8.8	3.7	47.6	23.1	0.0	0.3	0.2	1.2	0.5
Other illness	100%	15.2	3.0	9.8	4.9	4.4	59.0	0.3	0.7	0.8	1.4	0.5
Suicide	100%	0.7	0.2	0.6	0.1	0.6	2.5	92.5	0.5	0.6	1.4	0.4
Homicide	100%	3.2	0.0	2.2	0.0	0.7	3.2	0.0	86.0	0.0	1.8	2.9
Drug/alcohol intoxication	100%	6.5	0.0	0.8	0.5	1.0	10.3	3.0	0.8	65.1	1.0	11.1
Accident	100%	3.3	0.0	2.8	0.6	0.6	9.9	1.7	2.8	10.5	64.1	3.9
Other/unknown[b]	100%	21.0	1.8	11.4	5.7	6.6	36.2	5.7	0.4	5.2	3.9	2.2

Note: Includes only records with a known underlying cause of death in both the DCRP and NDI (n=15,963).

[a]Includes illnesses such as cerebrovascular disease, diabetes, and other nonleading natural causes of death.

[b]Includes cases where cause of death was provided but the manner was not known, cases where an autopsy was inconclusive, and cases where the information listed was a symptom rather than a cause or manner of death.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

The NDI attributed 18% of the 3,917 DCRP deaths due to heart disease (n=722) to some undefined illness categories, 4% to respiratory disease (n=148), and 4% to cancer (n=146). In total, the NDI attributed 3,553 deaths to heart disease, 9% less than the DCRP. The NDI also attributed 14% of DCRP homicides to other causes. Specifically, 9% were classified as natural deaths in the NDI—heart disease (3%), cancer (2%), respiratory or other illness (4%)—and the remainder were classified as accidental or other.

Close to half (49%) of deaths due to liver disease in the DCRP were attributed to other illness in the NDI. This included liver-related deaths caused by the hepatitis virus because of its infectious nature. The NDI also attributed 11% of DCRP deaths from respiratory disease to heart disease, 9% to cancer, and 23% to other illness.

When the NDI classified a record differently from the DCRP, it tended to place it in the same broad category of COD as in the DCRP (either illness or unnatural death). In total, 68% of the illness deaths in the DCRP were classified into the same major category of death in the NDI (table 8). Thirty percent of the illness deaths were attributed to a different illness category, while only 2% were reported as unnatural deaths. Eighty-six percent of the unnatural deaths in the DCRP received the same classification in the NDI, 8% were attributed to illness deaths, and 4% to other causes.

Assessing the sources of disagreement between the DCRP and NDI

Differences in the COD classification between the DCRP and NDI are due to several factors. Mismatched records are the first factor to consider. Two records may not be true matches even if they have the same identifiers, such as last name and birth and death dates. It is not possible to accurately distinguish true matches from false matches, but the NDI status code can be used as a proxy. As previously discussed, the NDI matching process assigned a status code of 1—indicating a high probabilistic score—to 82% of the matches, and a code of 0—indicating a low probabilistic score—to 18% of matches. The congruence in underlying COD between the DCRP and NDI should be higher for matches with a status code of 1 than for those with a status code of 0.

The second factor is data source. The information on the death certificate may be different from that on the medical examiner's report, or the text may be the same, but the case was assigned different ICD-10 codes by NDI and DCRP nosologists. This explanation was examined by directly comparing the sets of ICD codes from the DCRP and NDI for matched records.

The two datasets use different methods to identify the underlying COD when decedents have multiple health conditions. Specifically, the NDI uses ACME to identify the underlying COD from the list of ICD-coded conditions. When using ACME, shuffling the order in which the codes are listed can change the underlying COD. The DCRP instructs respondents to report a final COD as determined by a medical examiner. For multiple-cause deaths in the DCRP, the most common cause is designated as the underlying COD. Due to such procedural differences, even records with an identical set of ICD-10 codes may be assigned to different underlying COD categories in the two datasets. To understand how much this factor affects the congruence of the underlying COD between the DCRP and NDI, agreement rates were calculated separately for records with an identical set of ICD codes in the two datasets and those with different ICD codes. Agreement rates were also calculated by the number of ICD codes. The agreement rates should be higher for records with an identical set of ICD codes and for those with fewer ICD codes.

TABLE 8

Percent of cases in the DCRP and NDI that match on cause of death, 2007–10

DCRP cause of death	Same category as DCRP	Different category from DCRP		
		Illness	Unnatural	Other/ unknown[a]
All causes	69.8%	27.3%	2.2%	0.7%
Illness[b]	67.7%	30.3%	1.7%	0.3%
Heart disease	70.1	28.1	1.6	0.2
AIDS-related	78.8	20.2	1.1	0.0
Cancer	87.4	11.8	0.8	0.1
Liver disease	38.3	60.1	1.2	0.3
Respiratory disease	47.6	50.2	1.7	0.5
Other illness	59.0	37.3	3.2	0.5
Unnatural	85.8%	8.1%	3.6%	2.5%
Suicide	92.5	4.6	2.5	0.4
Homicide	86.0	9.3	1.8	2.9
Drug/alcohol intoxication	65.1	19.1	4.8	11.1
Accident	64.1	17.1	14.9	3.9

[a]Includes cases where cause of death was provided but the manner was not known, cases where an autopsy was inconclusive, and cases where the information listed was a symptom rather than a cause or manner of death.

[b]Includes illnesses such as cerebrovascular disease, diabetes, and other nonleading natural causes of death.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

The congruence of the underlying COD between the DCRP and the NDI was compared by status code and other factors (table 9). The kappa statistic of 0.46 and the overall agreement rate of 56% for records with a status code of 0 (i.e., matches with lower probabilistic match scores) were significantly lower than the kappa statistic of 0.66 and 72% for the corresponding statistics for those with a status code of 1 (i.e., those with higher probabilistic match scores). Except for accidental and other or unknown COD, deaths in each COD category by DCRP had a higher agreement rate for matches with a status code of 1 than for matches with a status code of 0. For example, 83% of the AIDS-related deaths in the DCRP with a status code of 1 were also classified as AIDS-related deaths in the NDI, compared to 57% of those with a status code of 0.

The agreement rates were also compared by the number of conditions (i.e., single-cause compared to multiple-cause deaths), facility type, sex, and death year. In this context, single-cause deaths were those that had only one ICD-10 code in both datasets, and multiple-cause deaths had more than one ICD-10 code in either the NDI or DCRP. Because no ambiguity was involved in identifying the underlying COD for single-cause deaths, single-cause deaths had higher intercoder reliability, with a kappa statistic of 0.77 and an overall agreement rate of 83%, compared to a kappa statistic of 0.52 and an overall agreement rate of 63% for multiple-cause illness deaths. The difference in agreement rates was also substantial within each illness category. For example, 73% of deaths attributed only to respiratory disease in the DCRP were also categorized as respiratory disease in the NDI, compared to 42% of the deaths attributed to respiratory disease in the DCRP that also had other causes. Because the DCRP did not assign ICD-10 codes to unnatural deaths, this analysis was restricted to DCRP illness deaths only.

Jail deaths had better agreement than prison deaths, with a kappa statistic of 0.68 (compared to 0.61 for prison deaths) and an overall agreement rate of 74% (compared to 69%). This was likely because nearly all jail deaths in the study group were sent for autopsy (95%), while about two-thirds (65%) of prisoner deaths were autopsied.

DCRP–NDI agreement improved over time. Within 4 years, the kappa statistic increased from 0.60 in 2007 to 0.68 in 2010 and the overall agreement rate increased from 67% in 2007 to 74% in 2010.

To further explore how DCRP–NDI agreement varies by decedent demographics and circumstances of death, a logistic regression was run with the 15,293 death records with COD information in both datasets. The dependent variable is binary, with 1 indicating that the record received the same COD classification in the DCRP and the NDI, and 0 indicating otherwise. The independent variables include facility type, sex, race or Hispanic origin, age group of the deceased, year of death, DCRP COD category, match quality (or NDI's status code), and the number of conditions (i.e., single compared to multiple).

TABLE 9

DCRP–NDI cause of death (COD) agreement, by status code, COD count, and facility type, 2007–10

		Status code[a]		Number of CODs		Facility type		Year			
	All	0	1	Single	Multiple	State prison	Local jail	2007	2008	2009	2010
Kappa[b]	0.64	0.46	0.66	0.77	0.52	0.61	0.68	0.60	0.61	0.66	0.68
Agreement rate											
All causes	69.8%	56.0%	71.8%	:	:	68.7%	74.1%	66.8%	67.6%	71.6%	73.5%
Illness	67.7%	56.2%	69.5%	83.0%	62.6%	67.6%	68.5%	64.5%	66.6%	69.5%	70.6%
Heart disease	70.1	60.7	71.4	84.7	66.7	68.9	74.9	68.9	70.2	70.6	70.6
AIDS-related	78.8	57.1	82.5	92.5	65.8	79.5	76.5	78.2	76.9	79.3	81.3
Cancer	87.4	72.0	89.8	94.3	83.3	87.6	81.2	85.5	88.3	86.9	88.5
Liver disease	38.3	23.8	40.5	63.2	33.2	37.3	48.0	36.4	35.5	40.8	40.9
Respiratory disease	47.6	36.5	49.0	72.5	42.3	46.2	57.5	45.3	48.3	49.8	47.2
Other illness[c]	59.0	50.1	60.4	65.3	57.2	58.3	62.1	54.7	56.0	62.1	65.1
Unnatural	85.8%	61.2%	88.4%	:	:	84.6%	86.9%	85.8%	84.8%	84.8%	87.4%
Suicide	92.5	64.7	95.3	:	:	91.0	93.6	92.5	92.3	92.6	92.4
Homicide	86.0	55.6	90.5	:	:	88.7	77.6	87.1	86.0	83.6	87.2
Drug/alcohol intoxication	65.1	46.9	66.7	:	:	64.8	65.3	65.1	61.5	65.4	68.5
Accident	64.1	65.0	64.0	:	:	64.5	63.4	59.5	66.7	56.6	73.1
Other/unknown[d]	2.2	0.0	2.5	:	:	1.6	2.8	2.5	1.0	0.0	13.3

Note: Includes only records with a known cause of death in both the DCRP and NDI (n=15,963).

: Not calculated. Unnatural deaths in the DCRP were not assigned International Classification of Diseases (ICD-10) codes.

[a]The NDI assigns a status code for each matched record. A status code of 1 indicates a high probabilistic match score, and a code of 0 indicates otherwise.

[b]Kappa is a measure of interrater agreement for qualitative items. It is a more robust measure than percent for agreement, as it takes into account the agreement occurring by chance.

[c]Includes illnesses such as cerebrovascular disease, diabetes, and other nonleading natural causes of death.

[d]Includes cases where cause of death was provided but the manner was not known, cases where an autopsy was inconclusive, and cases where the information listed was a symptom rather than a cause or manner of death.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

The regression results confirm the relationships observed in the previous analysis (table 10). Records with a status code of 1 were more likely to receive the same categorization for underlying COD in both datasets than those with a status code of 0, with an odds ratio of 2.2. This effect was statistically significant. Likewise, single-cause deaths were more likely to be consistently classified than multiple-cause deaths, with an odds ratio of 2.05.

Intercoder reliability also varied across COD categories. Deaths due to suicide, homicide, cancer, and AIDS-related problems had higher agreement rates than deaths by heart disease, while liver disease, respiratory disease, and drug and alcohol intoxication had lower agreement rates. Furthermore, deaths that occurred in jail had higher agreement rates than deaths that occurred in prison, with an odds ratio of 1.21. The sex of the decedent made no difference on agreement rate. Deaths of black decedents had lower agreement rates than those of white decedents, with an odds ratio of 0.91. There were no statistical differences among Hispanics, whites, and persons of other races. Lastly, the agreement rates in 2009 and 2010 were significantly higher than in 2007.

To further understand the source of mismatch between the underlying COD, detailed ICD-10 codes from the DCRP and the NDI were analyzed for matched records. This analysis was limited to 12,131 records with ICD codes from both records. Thirty percent of the cases were excluded either because they were classified as unnatural deaths, and therefore not assigned ICD-10 codes in the DCRP, or because no matches were found

TABLE 10

Logistic regression of DCRP–NDI cause of death agreement, 2007–2010

Dependent variable: Cause of death in agreement[a]	Odds ratio	Coefficient	Standard error
Status code = 1[b]	2.20*	0.12*	0.120
Single-cause death (ref = multiple-cause death)	2.05*	0.11*	0.106
DCRP cause of death category (ref = heart disease)			
AIDS-related	1.39*	0.17*	0.170
Cancer	2.82*	0.18*	0.179
Liver disease	0.26*	0.02*	0.018
Respiratory disease	0.38*	0.03*	0.028
Other illness[c]	0.61*	0.03*	0.032
Suicide	5.29*	0.52*	0.521
Homicide	2.94*	0.52*	0.522
Drug/alcohol intoxication	0.76*	0.09*	0.087
Accident	0.80	0.13	0.130
Other/unknown[d]	0.01*	0.00*	0.004
Local jail (ref = prison)	1.21*	0.06*	0.063
Female (ref = male)	0.98	0.08	0.081
Race/Hispanic origin (ref = white)			
Black/African American[e]	0.91*	0.04*	0.039
Hispanic/Latino	0.90	0.06	0.056
Other[f]	0.83	0.12	0.117
Death year (ref = 2007)			
2008	1.07	0.06	0.057
2009	1.16*	0.06*	0.062
2010	1.23*	0.07*	0.067
Constant	0.96	0.07	0.068
Observations	15,923		

Note: The reference group is the baseline used in comparisons. For example, an effect of 2.05 for single-cause deaths means that the odds of having a cause of death in agreement between the DCRP and the NDI were twice as high for single-cause deaths than for the reference group multiple-cause deaths.

*Indicates statistical significance at 0.05 level.

[a]The dependent variable is an indicator that equals 1 when DCRP and NDI major categories of death are in agreement and 0 when they are not.

[b]The NDI assigns a status code for each matched record. A status code of 1 indicates a high probabilistic match score, and a code of 0 indicates otherwise.

[c]Includes illnesses such as cerebrovascular disease, diabetes, and other nonleading natural causes of death.

[d]Includes cases where cause of death was provided but the manner was not known, cases where an autopsy was inconclusive, and cases where the information listed was a symptom rather than a cause or manner of death.

[e]Excludes persons of Hispanic or Latino origin.

[f]Includes American Indian or Alaska Native; Asian, Native Hawaiian, or Other Pacific Islander; and persons of two or more races.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

in the NDI. The 12,131 records were classified into 5 groups based on the relationship of their ICD codes reported in the DCRP and NDI (table 11).

Twenty percent of the cases (2,402) received the same set of ICD codes in the DCRP and NDI. In 3,371 cases (28%), the DCRP ICD codes were a subset of the NDI codes. That is, all DCRP codes were also reported on the matching NDI record, but the latter also had one or more codes not reported in the DCRP. The NDI codes were a subset of the DCRP codes in 342 cases (3%). Because the DCRP provides up to 5 ICD codes for each illness death, while the NDI provides up to 14 codes for each death in this dataset, it is more likely that the COD codes assigned by the DCRP were a subset of those assigned by the NDI than the reverse. In 1,129 cases (9%), the ICD-10 codes from the two datasets were different with some overlap. In 4,887 cases (40%), the ICD-10 codes from the two datasets were different with no overlap. The proportions of the last two groups indicate substantial mismatch in the all-cause ICD codes between the two datasets.

Ninety-four percent of records with the same set of ICD-10 codes in the DCRP and the NDI were assigned to the same major category of death by the two datasets in the categorization of the underlying COD (table 12). The agreement rate was 78% for records with DCRP codes being a subset of NDI codes, and 72% for records with NDI codes being a subset of DCRP codes. The agreement rate was lower for records with different sets of ICD codes with overlap (62%) and without overlap (52%).

Six percent of the cases with identical ICD-10 codes in the DCRP and the NDI ended up in different major categories of underlying COD. This could be due to the order in which the ICD-10 codes were entered into the system, as order can affect the final COD determination in ACME. For cases with a single identical ICD-10 code in both datasets, the agreement rate was almost 100%. As the number of ICD-10 codes increased to 5, the agreement rate dropped to 57%. The same pattern was observed for groups B and C. As the number of ICD-10 codes in the DCRP increased, the agreement rate dropped from 79% to 58% for group B, and from 84% to 57% for group C. For groups D and E, where DCRP and NDI codes were different with or without overlapping, the agreement rate was relatively low, regardless of the number of ICD-10 codes in the DCRP.

TABLE 11
Comparison of DCRP and NDI ICD-10 codes, 2007–10

Agreement of DCRP and NDI ICD-10 codes	Number	Percent
Total	12,131	100%
A. Identical codes in the DCRP and NDI[a]	2,402	19.8
B. DCRP codes were a subset of NDI codes[b]	3,371	27.8
C. NDI codes were a subset of DCRP codes[c]	342	2.8
D. NDI and DCRP codes overlapped[d]	1,129	9.3
E. No overlap between DCRP and NDI codes[e]	4,887	40.3

Note: Includes only matched records that had International Classification of Diseases (ICD-10) codes in both the DCRP and NDI datasets (n=12,131). The DCRP provides up to 5 ICD codes for illness deaths, while the NDI provides up to 14 codes for each matched case.

[a]This group received an identical set of ICD codes from the DCRP and NDI.

[b]This group received more ICD codes from the NDI than from the DCRP, and the DCRP codes were a subset of the NDI codes.

[c]This group received more ICD codes from the DCRP than from the NDI, and the NDI codes were a subset of the DCRP codes.

[d]The codes differed between DCRP record and its NDI match, but there was some overlap in the reported codes.

[e]The codes in the DCRP record and the NDI matched record were different and had no overlap.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

TABLE 12
Percent of records with same underlying cause of death, by International Classification of Diseases (ICD-10) codes, 2007–10

		Number of ICD-10 codes in the DCRP				
Agreement of DCRP and NDI ICD-10 codes	Total	One	Two	Three	Four	Five
Total	69.0%	70.6%	68.2%	61.7%	53.4%	56.7%
A. Identical codes in the DCRP and NDI[a]	93.5	99.7	78.4	76.3	65.2	57.1
B. DCRP codes were a subset of NDI codes[b]	78.1	78.7	76.3	69.0	58.3	~
C. NDI codes were a subset of DCRP codes[c]	71.9	~	83.6	60.2	48.8	56.5
D. NDI and DCRP codes overlapped[d]	61.9	~	65.4	61.5	52.0	58.3
E. No overlap between DCRP and NDI codes[e]	52.0	52.1	53.2	46.7	43.4	47.6

Note: Includes only matched records that had International Classification of Diseases (ICD-10) codes in both the DCRP and NDI datasets (n=12,131). The DCRP provides up to 5 ICD codes for illness deaths, while the NDI provides up to 14 codes for each matched case.

~Not applicable.

[a]This group received an identical set of ICD codes from the DCRP and NDI.

[b]This group received more ICD codes from the NDI than from the DCRP, and the DCRP codes were a subset of the NDI codes.

[c]This group received more ICD codes from the DCRP than from the NDI, and the NDI codes were a subset of the DCRP codes.

[d]The codes differed between DCRP record and its NDI match, but there was some overlap in the reported codes.

[e]The codes in the DCRP record and the NDI matched record were different and had no overlap.

Sources: Bureau of Justice Statistics, Deaths in Custody Reporting Program (DCRP), 2007–2010; and National Center for Health Statistics, National Death Index (NDI), 2007–2010.

These results confirm that disagreement in the underlying COD between the DCRP and the NDI was due partly to differences in the source data and partly to the different procedures for identifying the underlying COD. Forty percent of the cases received completely different ICD codes in the DCRP and the NDI. This group had a much lower agreement rate in underlying COD than records that received identical or similar ICD codes in the two datasets. Even among records that received the same ICD codes, the agreement rate declined as the number of conditions increased, suggesting that the different procedures for identifying the underlying COD in the two datasets can lead to different classifications of major categories of deaths.

Summary and implications

This study evaluated the feasibility of linking the DCRP data to the NDI database and assessed whether the DCRP was accurately capturing and measuring COD, whether the NDI could provide COD for DCRP cases missing this data, and whether the NDI could provide additional detail to illness deaths. Matches were found for 94% of the death records sent to the NDI, and 82% of the matches were considered a true match by the NDI matching program. Notably, the percentage of cases that were unable to be matched declined over time, from 7.3% in 2007 to 5.8% in 2010. To further improve the matching effectiveness of NDI search, BJS may consider the following:

- Create and use a unique subject ID variable for DCRP records sent to the NDI. A subject ID would allow the study to include persons who share the same last name, date of birth, and date of death as unique individuals with shared traits. Without the identifier, the distinction between unique persons and potential duplicate records could not be made, and these cases were excluded.
- Send multiple records with all known aliases to the NDI to improve matching. All aliases for one individual should be identified by the same subject ID. This would be applicable if the data providers included information about aliases by which an inmate is known (e.g., Sammy or Sam). If more than one match was returned, the match with the highest probabilistic score could then be used as the true match.
- Use decedents' SSNs in matching. The DCRP does not collect SSNs; however, BJS may be able to obtain them for decedents by matching DCRP records to other BJS administrative databases that include them.

The NDI linkage provides reliable, accurate, and comprehensive information of COD on jail and prison inmates. NDI data are based on death certificates and processed with standard National Center for Health Statistics (NCHS) codes. The NDI uses a reproducible automated algorithm (i.e., ACME) to identify the underlying COD. It allows comparisons with national data on noninstitutionalized populations because the same algorithms are used for ranking COD, determining underlying COD, and rolling up the CODs into broad categories of death. These algorithms are stable over time and systematic, and they are easily understood by public health and medical audiences.

The NDI provided more than twice as many ICD-10 codes per subject than did the DCRP (2.69 for the NDI compared to 1.06 for the DCRP). This was likely a result of the way CODs were recorded in death certificates. BJS may consider adopting the version found in standard death certificates.

With more accurate and comprehensive COD data from the NDI, BJS may produce statistics on specific diseases in correctional populations. Prison administrators and correctional health professionals may use detailed NDI trend data to plan and budget for treating chronic health conditions among long-term inmates. For example, hepatitis C has received a lot of attention in correctional health circles recently, in part because new pharmacotherapy to treat hepatitis C is expensive and the older standard course of treatment is complicated and hard to implement among inmates.[10] National statistics could be used by administrators and health care professionals to justify the funding and resources needed to mitigate the high rates of hepatitis C found in some facilities.

NDI search is a viable solution if BJS wants to reduce missing COD data or improve data quality. NDI search provided COD information for 90% of the cases with missing data in the DCRP. It takes approximately 12 months after the end of the calendar year for deaths to be included in the NDI database, which is 3 months longer than the DCRP. However, relying exclusively on the NDI for COD is not recommended because a subset of DCRP cases (6% in 2007–10) cannot be matched with NDI data, and some matches may not be correct.

[10]Keller, J. E. (2014). Hepatitis C treatment: Between a rock and a hard place. *Correct Care, 28*(2), 8-9, 20. http://www.ncchc.org/ filebin/ CorrectCare/28-2.pdf; and *Mortality in Local Jails and State Prisons, 2000–2013 - Statistical Tables*, NCJ 248756, BJS web, August 2015.

At the individual level, 70% of the records were classified into the same major category of death by the two datasets. The discrepancies were due to several factors: (1) the matched records may not be true matches, (2) most cases were not assigned the same set of ICD-10 codes by the two datasets, and (3) the process for identifying the underlying COD from multiple ICD codes was different between the two datasets such that even records with identical ICD codes may be classified to different major categories of death. Deaths with more ICD-10 codes were less likely to be categorized as the same underlying COD in the two datasets.

BJS publishes annual statistical tables on the distribution of COD over 11 major categories. When detailed NDI data were rolled into the same categories, the overall distributions from the two datasets were similar. In other words, the current DCRP method for obtaining COD produces aggregate statistics that are roughly comparable to what may be produced using NDI data.

Overall, the NDI linkage provides significant improvement in the quality of COD data in the DCRP. BJS will report COD statistics with increased confidence when DCRP COD data align with NDI data.

Methodology

Data source: Deaths in Custody Reporting Program

The Deaths in Custody Reporting Program (DCRP) is an annual Bureau of Justice Statistics (BJS) data collection. The DCRP collects national, state, and incident-level data on persons who died while in the physical custody of the 50 state departments of corrections or the approximately 2,800 local adult jail jurisdictions nationwide. The DCRP began in 2000 under the Death in Custody Reporting Act of 2000 (P.L. 106-297), and it is the only national statistical collection to obtain comprehensive information about deaths in adult correctional facilities. BJS uses DCRP data to track national trends in the number and causes (or manners) of deaths occurring in state prison or local jail custody.

The DCRP collects data about the characteristics of the decedents as well as circumstances surrounding the death, including the cause, time and location where the death occurred, and information on whether an autopsy was conducted and the availability of results to the respondent. Data on executions are excluded from this report but are accessible on the BJS website along with the DCRP mortality data.

The DCRP data collection instruments are administered annually to both state prisons and local jails. In addition to individual death records, respondents provide an aggregate count of the number of deaths that occurred during the referenced calendar year.

BJS obtains a separate report describing the decedent's characteristics and the circumstances surrounding the death for each death that occurred in a state prison or local jail. State prison and local jail respondents can submit individual records on decedents at any time during a collection cycle through a BJS web-based collection system.

Data source: The National Death Index

The National Death Index (NDI) is a centralized database of death certificates filed by state vital statistics offices. The NDI was established by the National Center for Health Statistics (NCHS) established the NDI as a mortality register for epidemiologists and other professional medical investigators and death researchers. Like the DCRP, the NDI is available for statistical purposes only, which precludes its use for legal, administrative, private or genological purposes.

The NDI is considered the "gold standard" for death identification because both the source, i.e., death certificates, and coverage is better than other mortality death files, such as the Social Security Administration's Death Master File.[11]

Nonresponse

All state department of corrections and more than 99% of jails participated for data years 2007 through 2010. However, there is varying degree of item nonresponse, particularly for the cause of death item. BJS changed data collection agents in 2009. As a result, the follow-up period for data year 2008 was cut short, and 21% of jail deaths and 3% of prison deaths were missing cause of death information. Aside from data year 2008, missingness affects a small percentage of cases—less than 1% of prison and 3% of jail deaths—annually.

Reported statistics

The kappa statistic was first proposed by Cohen.[12] It measures intercoder reliability while taking into account agreement that is expected purely by chance—that is, it indicates the level of agreement achieved beyond chance. For reliability between two raters (i.e., the DCRP and NDI), kappa is calculated as (po – pe) / (1 – pe) , where po is the observed proportionate agreement among raters and pe is the expected proportionate agreement by chance. When the raters are in complete agreement, kappa is 1. Landis and Koch suggested characterizing values less than 0 as indicating no agreement, 0 to 0.20 as slight, 0.21 to 0.40 as fair, 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial, and 0.81 to 1 as almost perfect agreement.[13]

[11]Fillenbaum, G.G., Burchett, B.M. & Blazer, D.G. (2009). Identifying a National Death Index Match. *American Journal of Epidemiology, 170*(4), 515–518.

[12]Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

[13]Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

The Bureau of Justice Statistics of the U.S. Department of Justice is the principal federal agency responsible for measuring crime, criminal victimization, criminal offenders, victims of crime, correlates of crime, and the operation of criminal and civil justice systems at the federal, state, tribal, and local levels. BJS collects, analyzes, and disseminates reliable and valid statistics on crime and justice systems in the United States, supports improvements to state and local criminal justice information systems, and participates with national and international organizations to develop and recommend national standards for justice statistics. Jeri M. Mulrow is acting director.

This report was written by Zhen Zeng, Margaret Noonan, and E. Ann Carson, BJS Statisticians; Ingrid Binswanger, BJS Visiting Fellow; Patrick Blatchford, Colorado School of Public Health; and Hope Smiley-McDonald, and Chris Ellis, RTI International. Anastasios Tsoutis and Emmaline Mitchell verified the report.

Lynne McConnell, Morgan Young, and Jill Thomas edited the report. Tina Dorsey and Barbara Quinn produced the report.

April 2016, NCJ 249568

BJS

Office of Justice Programs
Innovation • Partnerships • Safer Neighborhoods
www.ojp.usdoj.gov