Methodology Document

US Territory Mortality Statistics 2023

Drug-Related Deaths, Overdose Deaths, and Suicide Deaths
for Guam, Puerto Rico, Virgin Islands, American Samoa,
and Northern Mariana Islands

Prepared: December 10, 2025

Table of Contents

1. Overview and Request

This document describes the methodology used to calculate mortality statistics for US territories using CDC Multiple Cause of Death Public Use Files. The analysis was performed to create fact sheets for the following US territories:

- Guam (GU)
- Puerto Rico (PR)
- Virgin Islands (VI)
- American Samoa (AS)
- Northern Mariana Islands (MP)

The requested statistics included: (1) Drug-related deaths, (2) Overdose-related deaths, and (3) Suicide deaths for the year 2023.

2. Data Source

The data was obtained from the CDC National Center for Health Statistics (NCHS) Multiple Cause of Death Public Use Files. These files contain mortality data collected from death certificates filed in the United States.

2.1 Download Location

The data was downloaded from the CDC FTP server:

```
URL:  https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/DVS/mortality/
File: mort2023ps.zip
Unzipped file: VS23MORT.DPSMCPUB_r20241030
```

> **Note:** The CDC provides separate files for US states and US territories. The file used (mort2023ps.zip) contains data specifically for US possessions/territories.

2.2 File Characteristics

Characteristic	Value
Data Year	2023
File Format	Fixed-width ASCII text
Record Length	818 characters per line
Total Records (Territories)	36,476
Encoding	Latin-1

2.3 Documentation Reference

The file layout documentation was obtained from:

URL: https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/
DVS/mortality/2022-Mortality-Public-Use-File-Documentation.pdf

Note: The 2022 documentation applies to the 2023 file format as well.

3. File Format and Layout

The CDC mortality file uses a fixed-width format where each field occupies specific character positions. The following table shows the key fields used in this analysis:

3.1 Key Field Positions

Field Name	Position	Length	Description
State of Occurrence	21-22	2	FIPS state/territory code where death occurred
State of Residence	29-30	2	FIPS state/territory code of decedent's residence
Data Year	102-105	4	Year of death (e.g., 2023)
Manner of Death	107	1	Code indicating manner of death (1-7)
ICD-10 Underlying Cause	146-149	4	ICD-10 code for underlying cause of death

Important: Position numbers in the CDC documentation are 1-indexed. When using Python (0-indexed), subtract 1 from the starting position. For example, 'Position 21-22' in the documentation corresponds to `line[20:22]` in Python.

3.2 Sample Raw Data

Below is a sample record from the data file with key fields highlighted:

```
Sample line (first 160 characters):
                    11GU010  3GU  GU010            3               GU...
                    ^^       ^^
                    ||       ||
Position:          21-22    29-30
Field:             State    State
                   Occur    Resid
Value:              GU       GU
```

4. Field Definitions

4.1 Territory Codes

US territories are identified using 2-character FIPS codes. Statistics were calculated based on **State of Occurrence** (position 21-22), meaning deaths are attributed to the territory where the death occurred. This matches the CDC VSRR (Vital Statistics Rapid Release) methodology.

Foreign residents are excluded (resident_status = 4 at position 20) to match CDC WONDER methodology. Foreign residents are deaths occurring in the US where the decedent resided outside the US.

Code	Territory	Records in 2023 File
GU	Guam	1,183
PR	Puerto Rico	33,958
VI	Virgin Islands	750
AS	American Samoa	0 (no data)
MP	Northern Mariana Islands	232

4.2 Manner of Death Codes

The Manner of Death field (position 107) indicates how the death occurred:

Code	Description
1	Accident
2	Suicide
3	Homicide
4	Pending investigation
5	Could not determine
6	Self-Inflicted
7	Natural
Blank	Not specified

5. ICD-10 Code Classifications

The International Classification of Diseases, 10th Revision (ICD-10) codes are used to classify the underlying cause of death. The following sections describe how each statistic was defined using ICD-10 codes.

5.1 Drug Overdose Deaths

Drug overdose deaths are identified using ICD-10 codes for poisoning by drugs, medicaments, and biological substances. The following code ranges were used:

ICD-10 Range	Description	Intent
X40-X44	Accidental poisoning by drugs, medicaments and biological substances	Accidental
X60-X64	Intentional self-poisoning by drugs, medicaments and biological substances	Intentional (suicide)
Y10-Y14	Poisoning by drugs, medicaments and biological substances, undetermined intent	Undetermined

5.2 Drug-Related Deaths (Broader Definition)

Drug-related deaths include overdose deaths plus additional drug-induced causes. This broader category captures deaths where drugs played a significant role:

ICD-10 Range	Description
X40-X44, X60-X64, Y10-Y14	Drug overdose (as defined above)
F11-F16, F18-F19	Mental and behavioral disorders due to psychoactive substance use (excluding F10 alcohol and F17 tobacco)
T36-T50	Poisoning by drugs, medicaments and biological substances (therapeutic/accidental)

5.3 Suicide Deaths

Suicide deaths are identified using two methods:

Method 1: Manner of Death code = 2 (Suicide)
Method 2: ICD-10 codes X60-X84 (Intentional self-harm) or U03 (Terrorism involving suicide)

A death is counted as suicide if either condition is met.

ICD-10 Range	Description
X60-X84	Intentional self-harm (includes all methods: poisoning, hanging, firearm, etc.)
U03	Terrorism involving suicide

6. Data Processing Code

The following Python code was used to process the CDC mortality data file. The code reads the fixed-width file, extracts relevant fields, and calculates statistics for each territory.

6.1 Reading the Data File

```
# Open and read the fixed-width data file
DATA_FILE = "VS23MORT.DPSMCPUB_r20241030"

with open(DATA_FILE, 'r', encoding='latin-1') as f:
    for line in f:
        if len(line) < 150:
            continue  # Skip malformed lines

        # Extract fields (0-indexed, documentation is 1-indexed)
        state_occurrence = line[20:22].strip()  # Position 21-22
        state_residence = line[28:30].strip()   # Position 29-30
        manner_of_death = line[106:107]         # Position 107
        icd10_code = line[145:149].strip()      # Position 146-149

        # Process record...
```

6.2 Identifying Drug Overdose Deaths

```
def is_drug_overdose(icd10_code):
    """
    Check if ICD-10 code indicates drug overdose.
    Drug overdose codes:
    - X40-X44: Accidental poisoning by drugs
    - X60-X64: Intentional self-poisoning by drugs
    - Y10-Y14: Poisoning by drugs, undetermined intent
    """
    code = icd10_code.strip().upper()
    if len(code) < 3:
        return False

    # Check X40-X44 (accidental drug poisoning)
    if code.startswith('X4') and len(code) >= 3:
        digit = code[2]
        if digit in '01234':
            return True

    # Check X60-X64 (intentional self-poisoning by drugs)
    if code.startswith('X6') and len(code) >= 3:
        digit = code[2]
        if digit in '01234':
            return True

    # Check Y10-Y14 (undetermined intent drug poisoning)
    if code.startswith('Y1') and len(code) >= 3:
        digit = code[2]
        if digit in '01234':
            return True

    return False
```

6.3 Identifying Drug-Related Deaths

```
def is_drug_related(icd10_code):
    """
    Check if ICD-10 code indicates any drug-related death.
    Broader than overdose, includes:
    - Drug overdose (X40-X44, X60-X64, Y10-Y14)
    - Mental disorders due to drug use (F11-F16, F18-F19)
    - Drug toxicity and adverse effects (T36-T50)
    """
```

```
code = icd10_code.strip().upper()
if len(code) < 3:
    return False

# Drug overdose codes
if is_drug_overdose(icd10_code):
    return True

# Mental/behavioral disorders from drugs (F11-F19, excluding F10, F17)
if code.startswith('F1') and len(code) >= 3:
    digit = code[2]
    # Include F11-F16, F18-F19 (exclude F10=alcohol, F17=tobacco)
    if digit in '123456789' and digit != '0' and digit != '7':
        return True

# Drug poisoning/toxicity codes (T36-T50)
if code.startswith('T') and len(code) >= 3:
    try:
        num = int(code[1:3])
        if 36 <= num <= 50:
            return True
    except ValueError:
        pass

return False
```

6.4 Identifying Suicide Deaths

```
def is_suicide(manner_code, icd10_code):
    """
    Check if death is suicide.
    - Manner of death = 2 (Suicide)
    - OR ICD-10 codes X60-X84 (Intentional self-harm)
    - OR ICD-10 code U03 (Terrorism involving suicide)
    """
    # Check manner of death
    if manner_code == '2':
        return True

    # Check ICD-10 codes for intentional self-harm
    code = icd10_code.strip().upper()
    if len(code) < 3:
        return False

    # X60-X84: Intentional self-harm
    if code.startswith('X'):
        try:
            num = int(code[1:3])
            if 60 <= num <= 84:
                return True
        except ValueError:
            pass

    # U03: Terrorism involving suicide
    if code.startswith('U03'):
        return True

    return False
```

6.5 Main Processing Loop

```
from collections import defaultdict

# Territory codes
TERRITORIES = {'GU': 'Guam', 'PR': 'Puerto Rico', 'VI': 'Virgin Islands',
               'AS': 'American Samoa', 'MP': 'Northern Mariana Islands'}

# Initialize statistics
stats = defaultdict(lambda: {
    'total_deaths': 0,
    'suicide_deaths': 0,
    'drug_overdose_deaths': 0,
    'drug_related_deaths': 0,
})

# Process each record
with open(DATA_FILE, 'r', encoding='latin-1') as f:
    for line in f:
        if len(line) < 150:
            continue

        # Use State of Occurrence (position 21-22) to match CDC VSRR methodology
        state_occurrence = line[20:22].strip()
        manner_of_death = line[106:107]
        icd10_code = line[145:149].strip()

        if state_occurrence not in TERRITORIES:
            continue

        stats[state_occurrence]['total_deaths'] += 1

        if is_suicide(manner_of_death, icd10_code):
            stats[state_occurrence]['suicide_deaths'] += 1

        if is_drug_overdose(icd10_code):
            stats[state_occurrence]['drug_overdose_deaths'] += 1

        if is_drug_related(icd10_code):
```

```
stats[state_occurrence]['drug_related_deaths'] += 1
```

7. Results

The following table presents the final mortality statistics for each US territory based on the 2023 CDC Multiple Cause of Death Public Use File:

7.1 Summary Statistics

Territory	Total Deaths	Suicide Deaths	Overdose Deaths	Drug-Related Deaths
Puerto Rico	34,290	236	786	2,412
Guam	1,193	31	34	229
Virgin Islands	758	10	5	23
American Samoa	N/A*	N/A	N/A	N/A
Northern Mariana Islands	235	4	0	25

*American Samoa did not report mortality data for 2023 (see CDC NVSR Vol 74, No 8: https://www.cdc.gov/nchs/data/nvsr/nvsr74/nvsr-74-08.pdf).

7.2 Validation: Public-Use Files vs CDC WONDER

To validate our methodology, we compared **national US overdose death counts** from both data sources using identical ICD-10 codes (X40-X44, X60-X64, X85, Y10-Y14) as underlying or contributing cause.

CDC WONDER excludes foreign residents - deaths that occurred in the US where the decedent resided outside the US (resident_status = 4 in the public-use files). When we exclude foreign residents from the public-use files, the counts match CDC WONDER exactly:

Metric	CDC WONDER	Public-Use (ALL)	Public-Use (excl. foreign)
2023 Overdose Deaths	112,106	114,121 (+1.8%)	112,106 (exact match)
2023 Total Deaths	3,090,964	3,101,016 (+0.3%)	3,090,964 (exact match)

In 2023, there were 10,052 foreign resident deaths (0.32% of total) and 2,015 foreign resident overdose deaths (1.77% of overdoses). This analysis excludes foreign residents to match CDC WONDER methodology.

8. Appendix: Replication

The complete Python scripts used for data processing and fact sheet generation are included with this document:

• **process_mortality_data.py** - Main data processing script

• **create_report.py** - PDF summary report generator

• **create_methodology_document.py** - This methodology document generator

8.1 Data Files

```
Data Files:
  VS23MORT.DPSMCPUB_r20241030     (Raw mortality data from CDC)
  2023-Mortality-Public-Use-File-Documentation.pdf (File layout documentation)

Output:
  territory_mortality_summary_2023.csv
  US_Territory_Mortality_Statistics_2023.pdf
  Methodology_US_Territory_Mortality_Statistics.pdf
```

8.2 Replication Steps

To verify these results, download the CDC mortality data file and run the processing script:

```
# Download data from CDC
wget https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/DVS/mortality/mort2023ps.zip
unzip mort2023ps.zip

# Run processing script
python3 process_mortality_data.py
```

Date: December 10, 2025

Data Source: CDC NCHS Multiple Cause of Death Public Use Files, 2023