

**Division of Health Informatics and Surveillance (DHIS)  
Data Processing and Security Procedures for Nationally Notifiable Diseases Data  
Center for Surveillance, Epidemiology and Laboratory Services (CSELS)  
Centers for Disease Control and Prevention (CDC)  
Atlanta, GA 30333**

Last Revised: January 15, 2015  
Program Lead: Umed Ajani, DHIS, CSELS  
Phone: (404) 498-0258  
E-mail: uaa0@cdc.gov

## **Table of Contents**

- 1. Introduction**
- 2. Program description**
- 3. Nationally notifiable infectious diseases data collection**
- 4. Privacy Act and private information in data reported to DHIS**
- 5. Procedures for assuring data quality**
- 6. DHIS data use**
- 7. Procedures for maintaining security of information during data transmission and use**
- 8. Procedures for maintaining privacy of information during publication**
- 9. Procedures for data release (public use data sets)**
- 10. Procedures for maintaining privacy of information in release of public use data sets**
- 11. Sharing of DHIS nationally notifiable infectious disease data**
- 12. Human subjects protection**

### **1. Introduction**

This document describes the procedures for the Division of Health Informatics and Surveillance (DHIS), Center for Surveillance, Epidemiology and Laboratory Services (CSELS), for receiving, securing, provisioning, publishing, and releasing nationally notifiable infectious diseases data received at CDC by information systems administered by DHIS. The procedures in this document apply only to the collection, security, provisioning, publishing and release of nationally notifiable infectious disease data by DHIS.

Surveillance programs in the National Center for Emerging and Zoonotic Infectious Diseases (NCEZID), National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention (NCHHSTP), National Center for Immunization and Respiratory Disease (NCIRD), and the Center for Global Health (CGH) have primary responsibility at CDC for surveillance of the infectious diseases and conditions covered by each Center. Programs within the Centers receive nationally notifiable infectious disease data from DHIS and use, release and/or share their programs' data according to guidance established by CDC, their Centers and programs (e.g., <http://www.cdc.gov/nchhstp/programintegration/Data-Security.htm/> and <http://www.cdc.gov/nchhstp/atlas/>).

In this document, the term “data sharing” refers to granting certain individuals or organizations

outside of CDC access to data that may contain personally identifiable health information (PII) after signing a formal data sharing agreement with a CDC data steward. Office of Management and Budget (OMB) Memorandum M-07-16 defines personally identifiable information as “information that can be used to distinguish or trace an individual’s identify either alone or when combined with other personal or identifying information that is linked or associated with the individual.” The term “data release” refers to the dissemination of public use data sets without PII in a format that minimizes the chance of identification. The term “data provisioning” refers to DHIS’s providing condition-specific data to the CDC programs responsible for national level surveillance, prevention and control for each condition.

## **2. Program Description**

DHIS’s receipt, securing, provisioning, publishing and releasing notifiable infectious disease data are components of the National Notifiable Diseases Surveillance System (NNDSS). DHIS also supports state, local and territorial health departments in collecting, managing, and analyzing their data and in submitting case notification data to CDC for NNDSS by providing funding, health information exchange standards and frameworks, electronic health information systems and technical support.

NNDSS is a nationwide collaboration that enables all levels of public health (local, state, territorial, and federal) to share health information required to monitor, control, and prevent the occurrence and spread of infectious and some noninfectious diseases and conditions. NNDSS is a multifaceted program that includes the surveillance system for collection, analysis, and sharing of health data and also policies, laws, electronic messaging standards, people, partners, information systems, processes, and resources at the local, state, territorial and national levels.

More information on NNDSS, including case definitions, is available at:  
[http://wwwn.cdc.gov/nationally notifiable infectious disease data/](http://wwwn.cdc.gov/nationally_notifiable_infectious_disease_data/)

## **3. Nationally Notifiable Infectious Diseases Data Collection**

Notifiable infectious disease surveillance begins at the level of local, state, and territorial health departments. Jurisdictional laws and regulations mandate reporting of cases of specified infectious and noninfectious conditions to health departments. The health departments work with healthcare providers, laboratories, hospitals, and other partners to obtain the information needed to monitor, control and prevent the occurrence and spread of these health conditions. Additionally, health departments notify CDC about the occurrence of certain conditions.

A number of different legislative and regulatory decisions have authorized federal collection of notifiable disease data and given responsibility for that data collection to CDC ([http://wwwn.cdc.gov/nationally notifiable infectious disease data/script/history.aspx/](http://wwwn.cdc.gov/nationally_notifiable_infectious_disease_data/script/history.aspx/)). In 1961, responsibility for the collection of data on nationally notifiable diseases and deaths in 122 U.S. cities was transferred from the National Office of Vital Statistics to CDC. CDC is responsible for the dissemination of nationally notifiable diseases’ information, as authorized by the Public Health Service Act (42 USC 241) of January 4, 2012.

CDC programs in NCEZID, NCHHSTP, NCIRD and CGH that are responsible for national surveillance, prevention and control of infectious conditions collaborate with the Council of State and Territorial Epidemiologists (CSTE) to determine which conditions reported to the jurisdictions are nationally notifiable, and the CDC programs, in collaboration with subject matter experts in CSTE and in jurisdictions, determine what data elements are included in national notifications. Core data elements, with common definitions, are collected for all of the infectious diseases submitted to CDC. The core data elements include the name of the condition; demographic data; administrative data such as the jurisdiction submitting the case; epidemiologic data; characteristics of the disease; and contact information for the person submitting the case to CDC. For many of the conditions, the public health department may also submit data elements which are specific to the condition. The disease-specific data elements vary by disease and may include signs and symptoms, additional diagnostic data, treatment data, vaccination history, laboratory tests and results, and risk factors.

Jurisdictions participating in NNDSS voluntarily submit infectious disease data to DHIS and may also submit some data directly to NCEZID, NCHHSTP, NCIRD, and CGH. The infectious conditions and data elements for which the DHIS information systems receive data have changed over the years in response to changes in the list of nationally notifiable conditions, changes in case definitions and data elements, and in response to efforts to coordinate data collection efforts with NCEZID, NCHHSTP, NCIRD, and CGH.

The transmission of data to DHIS is supported by several interconnected health information exchange frameworks and electronic information systems. Most case records are submitted electronically from already existing data-bases via automated electronic transfers through a secure network. Supplemental information may be received at DHIS by telephone, fax, mail, and email. Procedures for maintaining security of the data during transmission are described in section 7 below.

#### **4. Privacy Act and Personally Identifiable Information in Data Reported to DHIS**

The nationally notifiable infectious disease data received by DHIS contain sensitive personally identifiable health information (PII), which are subject to the Privacy Act. As noted above, Office of Management and Budget (OMB) Memorandum M-07-16 defines personally identifiable information as “information that can be used to distinguish or trace an individual’s identity either alone or when combined with other personal or identifying information that is linked or associated with the individual.” The Privacy Act is a Federal law that protects PII held by a federal agency in a system of records from which information is retrieved by an individual’s name, an identification number, or some other unique identifier assigned to the individual. The CDC Privacy Act System of Records Notice that covers NNDSS is 09-20-0136. This notice provides information to the public about the existence of the CDC research and surveillance systems covered by the notice, how data are used, how data are safeguarded, and how information may be disclosed (<http://www.cdc.gov/SORNnotice/09-20-0136.htm/>).

Some personal identifiers are submitted to DHIS for all nationally notifiable infectious conditions:

- a unique Case Notification ID for the specific disease or condition for a given patient,

- IDs for linking case notification data to local health records, and
- the date of birth (month, day, and year) of the person with the condition.

For some conditions, the name of the individual with the condition is submitted along with IDs for linking with additional local health records.

In addition, some combinations of submitted data elements, including date of birth, date of death (month/day/year), sex, ethnic group, race, state, county, and ZIP code could potentially be used to identify individuals.

The nationally notifiable infectious diseases electronic records received by DHIS are stored at DHIS indefinitely.

DHIS implements a number of safeguards to prevent the disclosure of PII and to prevent the identification of individuals when data are published, released or shared (Sections 7, 8, 10, and 12, below).

## **5. Procedures for Assuring Quality of Data**

On a weekly basis, DHIS staff, in collaboration with subject matter experts in NCEZID, NCHHSTP, NCIRD, and CGH, evaluate the information submitted and review case reports for missing data elements, for inconsistencies among data elements and for data values that are not consistent with the values identified in the Public Health Information Network (PHIN) Vocabulary Access and Distribution System (VADS) code sets. When potential errors are identified, DHIS staff works with the individuals at the state or local health department who report the data to DHIS to resolve errors. Only data that meet data quality criteria are published or released.

On an annual basis, DHIS staff works with the staff in state and territorial health departments and with staff in NCEZID, NCHHSTP, NCIRD, and CGH to finalize reports of cases for that year with local or county health departments and reconcile the data with reports previously sent to CDC throughout the year; collaborating with state health reporters, state epidemiologists, and CDC program offices to reconcile and verify counts and obtain approval from the appropriate chief epidemiologist from each state or territory before final counts are published. Only data that meet data quality criteria are published or released.

## **6. DHIS Data Use**

A major purpose of the DHIS-administered information system is to provide mechanisms by which CDC may receive nationally notifiable infectious disease data submitted by local, territorial and state health departments through one organization (DHIS), assess and assure the quality of the data and provide those data to the CDC infectious disease surveillance programs in NCEZID, NCHHSTP, NCIRD, and CGH that have responsibility for national surveillance, prevention and control efforts. Data are used by CDC infectious disease surveillance programs to monitor the occurrence of the conditions, identify populations or geographic areas at high risk, plan prevention and control programs and policies, allocate resources appropriately, and evaluate

the effectiveness of programs and policies. In addition, information is collected that allow infectious diseases programs to trace cases and their contacts and their travel histories, or other linkages necessary to describe and manage outbreaks or conduct public health follow-up to minimize the spread of disease.

Another major DHIS activity is to publish in the *MMWR* reports on notifiable conditions. Numbers of provisional cases of nationally notifiable infectious diseases submitted to CDC during the current week are included in the weekly *MMWR* morbidity report. Once each year, DHIS publishes an “early release” of finalized notifiable disease case counts for a given year. DHIS also publishes finalized notifiable diseases data in more detail in a later annual *Summary of Notifiable Diseases, United States*. The annual report is used to update annual tables published by the World Health Organization, the Pan American Health Organization, the U.S. Bureau of the Census, and the National Center for Health Statistics. DHIS provides the official source of statistics in the United States for nationally notifiable conditions and CDC is the sole repository for these national, population-based data.

DHIS also releases public use data sets with de-identified aggregated data tables in the same format as published in *MMWR*. Those data sets are made available CDC’s web-based query system known as CDC WONDER (<http://wonder.cdc.gov/>) and on Data.CDC.Gov (<https://data.cdc.gov/>) (Sections 9-10, below).

## **7. Procedures for Maintaining Security of Information during Data Transmission and Use**

The security of private information during electronic transmission to DHIS is maintained by technologies (computers and servers) that use national public health standards for messaging systems which provide security mechanisms for jurisdictions to use when submitting data. Most case records are encrypted and submitted to DHIS electronically from already existing databases via automated electronic transfers through a secure network. Electronic data are transmitted to and processed within the electronic information system platforms. The electronic data are treated in a secure manner consistent with the technical, administrative, and operational controls required by the Federal Information Security Management Act of 2002 (FISMA). These systems are also in compliance with more recent standards to protect information: the NIST Recommended Security Controls for Federal Information Systems and Organizations, Special Publication 800-53, Revised May 1, 2010. On occasion, when electronic transmission is not possible or when public health departments prefer, weekly case counts are provided by telephone, fax, mail, and email, primarily to meet weekly deadlines for publication in the *MMWR*. For these non-electronic transmissions, no identifiers are included, and safeguards are implemented to ensure that information is received only by the appropriate staff at DHIS. DHIS provides nationally notifiable infectious disease data to CDC programs through secure electronic platforms. The platforms are subject to CDC’s Certification and Accreditation process, in which these controls are examined and validated by the CDC’s Office of the Chief Information Security Officer.

All hard copy materials submitted to CDC are stored in locked cabinets in restricted access areas in buildings that require card key access.

PII information will not be disclosed unless otherwise compelled by law.

## 8. Procedures for Maintaining Privacy during Publication

During publication, privacy of the data is ensured through publication of only limited information about each case reported to DHIS. Most of the data elements, including individual identifiers are not published. Only counts and rates of each condition are published for individuals with the condition categorized by demographic characteristic (sex, age, race, and ethnicity categories), and geographic location (national, state, territorial or county level), as well as the time of the report to CDC (week, month or year), as described below.

In weekly MMWR reports, DHIS publishes the following provisional counts:

1. At the national level, for infrequently reported notifiable conditions, the number of cases reported during the current week, the cumulative count for the current year, the 5-year weekly case count average, the total case count total case counts for 5 previous years, and the names of the states and territories reporting cases during the current week along with the number of cases the state or territory is reporting.
2. For states, territories, regions, and the nation, for selected conditions, the number of cases reported during the current week, the median and maximum numbers of cases reported during the past 52 weeks, and cumulative year-to-date counts for the current and previous years.
3. Annual final national case counts and incidence rates by residency status (U.S. residents and non-residents).

No counts by sex, age, race, or ethnicity are published except for *H. influenzae*, for which counts by age group are provided.

In annual MMWR reports of final (“early release”) case counts, DHIS publishes the following:

1. For states and territories, regions and the nation, final annual notifiable condition case counts, by state, territory and nation.
2. Annual final national case counts and incidence rates by residency status (U.S. residents and non-residents).

No cross-tabulations are published.

Case counts as low as 1 and rates as low as 0.00 (less than 1 per 1,000) are published. However, if the total number of cases in a reporting jurisdiction for any condition is less than or equal to 25, then race-specific and ethnicity-specific information is not published.

In annual MMWR summary of notifiable diseases Part 1, DHIS publishes data tables with the following:

1. Nationally, final annual and monthly case counts for each condition.
2. For states and territories, regions, and the nation, final annual case counts for notifiable conditions by geographic location.

3. Annual final national case counts and incidence rates by the following age groups: <1 yr, 1-4 yrs, 5-14 yrs, 15-24 yrs, 25-39 yrs, 40-64 yrs, and >65 yrs.
4. Annual final national case counts and incidence rates by sex.
5. Annual final national case counts and incidence rates by race.
6. Annual final national case counts and incidence rates by ethnicity.
7. Annual final national case counts and incidence rates by residency status (U.S. residents and non-residents).

No cross-tabulations are published.

Case counts as low as 1 and rates as low as 0.00 (less than 1 per 1,000) are published. However, if the total number of cases in a reporting jurisdiction for any condition is less than or equal to 25, then race-specific and ethnicity-specific information is not published.

In annual MMWR summary of notifiable diseases Part 2, DHIS publishes the following:

1. Graphs with national level numbers and incidence rates for selected conditions.
2. Maps with state or county level case counts and incidence rates for selected conditions.
3. Graphs or maps with national level case counts and incidence rates by race-ethnicity, by gender, and by birthplace (US/non-US).

Case counts as low as 1 and rates as low as 0.00 (less than 1 per 1,000) are published.

In annual MMWR summary of notifiable diseases Part 3, DHIS publishes tables with the following:

1. National final annual case counts and rates by year, by condition.

No cross-tabulations are published.

Case counts as low as 1 and rates as low as 0.00 (less than 1 per 1,000) are published. However, if the total number of cases in a reporting jurisdiction for any condition is less than or equal to 25, then race-specific and ethnicity-specific information is not published.

## **9. Procedures for Data Release (Public Use Data Sets)**

DHIS releases downloadable, machine-readable public use data sets in two locations:

CDC WONDER (<http://wonder.cdc.gov/>)

HHS Data.gov (<https://data.cdc.gov/>)

The data tables are identical to those published in the weekly *MMWR* and in the *MMWR* annual summary part 1 as described in section 8 above. Privacy of the data is ensured through release of only limited information about each case reported to DHIS. Data tables are released only in the format of the published tables and cross-tabulations provide data at no more detailed level than is described in section 8 above.

Historical data tables with annual summary final case counts and rates and weekly provisional case counts for the years 2000 – 2014 are being made available, following the same policy as described above.

Data documentation for understanding the data elements in the tables is provided in the footnotes to the tables.

#### **10. Procedures for Maintaining Privacy of Information in Release of Public use Data Sets**

The procedures for protecting private information with the release of data are the same as those described for publication in section 8 and 9 above.

#### **11. Sharing of nationally notifiable infectious disease data**

DHIS does not share nationally notifiable infectious disease data with individuals or organizations outside of CDC. DHIS provides nationally notifiable infectious disease data only to the specific CDC program that is responsible for surveillance of that specific condition. When DHIS receives requests for data from individuals or organizations outside of CDC, DHIS notifies the requestor that the DHIS NNDSS program is unable to fill requests for specific datasets at this time and that some NNDSS data are available on CDC WONDER (<http://wonder.cdc.gov/>) and DATA.GOV (<https://www.data.gov/>). In addition, DHIS provides a link to the web site of the CDC program that is responsible for surveillance of the condition for which data are requested and notes the availability of data and a program point of contact at that site.

#### **12. Human Subjects Protection**

Although public use data sets that are released by DHIS may be used for research, no Institutional Review Board approval is required since the public use data sets are de-identified and only summary information is made available.