What Data Users Should Know About the
National Notifiable Diseases Surveillance System

The National Notifiable Diseases Surveillance System (NNDSS) database contains a rich source of infectious disease surveillance data for the United States. These data are important for monitoring trends of notifiable infectious diseases and for targeting research, prevention, and control efforts. The purpose of this document is to summarize important information data users should know about the National Notifiable Diseases Surveillance System data in order to facilitate proper analysis and interpretation of the data and to encourage data use.

Summary of Important Guidance to Users of NNDSS Data

1. NNDSS data are compilations of data voluntarily reported from U.S. States and Territories to CDC.
2. The national reporting status of diseases varies from year to year
3. The reporting status of diseases within each state is determined by state laws and regulations.
4. Annual state counts of NNDSS data are considered finalized when they are published as early release tables in the *Morbidity and Mortality Weekly Report* (*MMWR)*, which usually occurs in October or November each year for data derived primarily from the preceding year.
5. Before 1990, only summary electronic NNDSS data are available. Beginning in 1990, summary and case-specific electronic NNDSS data files are available for data release.
6. NNDSS data use criteria for classification and enumeration specified in national CDC/CSTE-sanctioned surveillance case definitions for infectious conditions of public health importance.
7. NNDSS case counts are likely incomplete, and therefore, these data are considered to represent a minimum number of cases.
8. Surveillance practices, policies, priorities, and resources vary from state to state. Therefore, one should use caution when making state-to-state comparisons of disease incidence.
9. Data on non-notifiable diseases/conditions may represent incidence data that are not population-based.
10. A substantial amount of time may elapse between the time a disease is made notifiable and the time surveillance data are available for that disease.
11. There may be differences between the case counts in state-specific surveillance databases and the case counts in the finalized NNDSS surveillance database.
12. There may be differences in case counts between the contents of the NNDSS surveillance database and the case counts printed in the *MMWR* weekly or *MMWR Summary of Notifiable Diseases*.
13. When calculating the incidence rate of a nationally notifiable disease, the denominator may need to be adjusted to exclude the population from states where the disease was not notifiable or data were unavailable.
14. NNDSS data pertaining to race and ethnicity may be incomplete and may have been

collected differently across the various notifiable diseases and across time.

15. States may batch report cases to the NNDSS database which may impact analyses of timeliness. In addition, some administrative procedures used in processing the NNDSS data may impact analyses of timeliness.
16. The NNDSS data may undergo limited logical error checking.
17. Varicella (chickenpox) and other non-notifiable diseases of public health importance may be transmitted by states to CDC for the national surveillance of these diseases in addition to the diseases designated as nationally notifiable.

Important Guidance to Users of NNDSS Data

1. NNDSS data are compilations of data voluntarily reported from U.S. States and Territories to CDC.

The NNDSS surveillance database includes data primarily reported from U.S. States and autonomous reporting entities, including New York City and Washington D.C. However, the database also includes limited data from the following U.S. Territories: American Samoa, the Commonwealth of Northern Mariana Islands, Guam, Puerto Rico, and the Virgin Islands. The reporting of nationally notifiable diseases to the Centers for Disease Control and Prevention (CDC) by states and territories (hereafter, collectively referred to as states) is voluntary. The Council of State and Territorial Epidemiologists (CSTE) recommends states also voluntarily report selected non-notifiable diseases such as varicella (chickenpox) infection to CDC.

2. **The national reporting status of diseases varies from year to year**

A notifiable disease is one for which regular, frequent, timely information on individual cases is considered necessary to prevent and control that disease. Each year a list of nationally notifiable diseases is agreed upon and maintained by CSTE and CDC. The list may vary by year as CSTE and CDC jointly make additions and/or deletions to the previous year's list. The most current list of nationally notifiable diseases is available on the Internet: http://www.cdc.gov/epo/dphsi/infdis.htm.

3. **The reporting status of diseases within each state is determined by state laws and regulations.**

Diseases that are considered nationally notifiable may or may not be designated by a given state as notifiable (reportable) in the state. States may use the national notifiable diseases list as well as other information, such as state-specific health priorities, to guide their determination of which conditions/diseases to make notifiable in their state. Thus, the list of state-specific notifiable diseases may vary across the states and, in a given state, the list may vary over time as well. Disease reporting is currently mandated by legislation or regulation only at the local or state level.

4. Annual state counts of NNDSS data are considered finalized when they are published as early release tables in the *Morbidity and Mortality Weekly Report* (*MMWR*), which usually occurs in October or November each year for data derived primarily from the preceding year.

Demographic data associated with NNDSS case count data, however, are not considered finalized until the *MMWR Summary of Notifiable Diseases* is published, which usually occurs at the end of each calendar year.

5. Before 1990, only summary electronic NNDSS data are available. Beginning in 1990, summary and case-specific electronic NNDSS data files are available for data release.

Before 1990, NNDSS data were transmitted to CDC via paper morbidity reporting

forms and these data were hand-tabulated for the *MMWR* weekly and *MMWR Summary of Notifiable Diseases*. Therefore, only electronic summary NNDSS data are available before 1990 and these data do not include demographic information. Summary data files can include name of the event (disease/condition of interest), case count, reporting state, reporting county, and the event month/year. Beginning in 1990, electronic case-specific data are available for release. However, some states still only do summary reporting for selected diseases. Please refer to the "Data Release Guidelines for the National Notifiable Diseases Surveillance System" for a description of the types of summary and case-specific data files that can be released by CDC.

6. NNDSS data use criteria for classification and enumeration specified in national CDC/CSTE-sanctioned surveillance case definitions for infectious conditions of public health importance.

CDC/CSTE-sanctioned national surveillance case definitions for infectious diseases were first published in October 1990. Historical and current versions of these case definitions are available on the CDC Home Page: http://www.cdc.gov/epo/dphsi/casedef/index.htm. These case definitions provide uniform criteria for classifying cases for purposes of public health surveillance. The intent of these case definitions is to improve the comparability of diseases reported from different geographic areas. Differences in the way states apply surveillance case definition criteria may result in differences in the way states classify and count cases.

7. NNDSS case counts are likely incomplete, and therefore, these data are considered to represent a minimum number of cases.

The NNDSS surveillance data likely represent a underestimate of the true number of cases of a given condition because of disease under-recognition and under-reporting. Under-recognition of a disease process can occur if a patient with mild or severe illness doesn't seek medical care or seeks medical care and the disease isn't accurately diagnosed by a health care provider. Under-reporting occurs when public health specialists aren't notified or otherwise aware of the occurrence of a disease that has importance to the public health community. Completeness of reporting is also influenced by the diagnostic facilities that are available to the clinician; the control measures that are in effect; the public awareness of a specific disease; and the interests, resources, and priorities of state and local officials responsible for disease control and public health surveillance. Factors such as changes in the case definitions for public health surveillance, the introduction of new diagnostic tests, or the discovery of new disease entities may cause changes in disease reporting that are independent of the true incidence of disease.

8. Surveillance practices, policies, priorities, and resources vary from state to state. Therefore, one should use caution when making state-to-state comparisons of disease incidence.

States establish rules and procedures for reporting data from within their jurisdictions to the NNDSS surveillance database; these rules and procedures may vary

by state. For example, some states may only report confirmed cases to CDC whereas others may report additional case classification categories such as probable, suspect, or unknown. Some of the CDC/CSTE surveillance case definitions include guidance on which case classification categories to report to CDC. EPO's Surveillance Systems Branch has not conducted a recent formalized assessment across NNDSS conditions to determine how this guidance is implemented.

Each state determines how to manage their priority health issues and resources. Therefore, each state determines whether a surveillance system for a given notifiable disease in their state should be active or passive in nature, population-based or case-based, and state-wide or localized to a specific region of the state. Because resources and priorities change over time, the attributes of a surveillance system for a given condition are also subject to change over time. The nature and characteristics of each state's surveillance system may impact the interpretation of an NNDSS-based data analysis.

9. Data on non-notifiable diseases/conditions may represent incidence data that are not population-based.

When a disease is considered non-notifiable in a given state, health care providers and laboratories are not mandated to report cases to public health authorities in that state. Therefore, data on non-notifiable diseases may not represent population-based incidence. When a state designates a disease as notifiable, this means that health care providers and/or laboratories are mandated to report cases to public health authorities and there is a formalized mechanism for identifying or receiving case reports, for conducting case follow-up, and initiating prevention activities. Therefore, data on notifiable diseases are more likely to represent incidence data that are population-based than are data on non-notifiable diseases.

10. A substantial amount of time may elapse between the time a disease is made notifiable and the time surveillance data are available for that disease.

Recommendations for modifications to the list of nationally notifiable diseases are made through the CSTE position statement process. Each year, position statements are introduced and adopted at the CSTE Annual Meeting. CSTE's adopted position statements are disseminated in finalized format in June or July of each year–these position statements may include recommendations for additions and deletions to the national notifiable diseases list. In January of the following year, CDC revises its official list of nationally notifiable diseases to coincide with CSTE position statement recommendations. A substantial period of time can elapse between the time CSTE recommends a new disease/condition for inclusion on the national notifiable diseases list and the time until a state can establish and fully operationalize a surveillance system specific to that disease/condition. Therefore, case ascertainment may be less complete in the years shortly after a disease becomes nationally notifiable than after the surveillance system has become fully operational.

11. There may be differences between the case counts in state-specific surveillance

databases and the case counts in the finalized NNDSS surveillance database.

The NNDSS database is not a real-time database. At each year's end the accumulated NNDSS data are verified, corrected, and finalized. After being finalized, the NNDSS data are no longer reconciled with state databases. Therefore, there can be differences between the case counts in the state-specific databases and CDC's NNDSS surveillance database; and, these differences may increase over time.

12. There may be differences in case counts between the contents of the NNDSS surveillance database and the case counts printed in the *MMWR* weekly or *MMWR Summary of Notifiable Diseases*.

The weekly NNDSS data are provisional and can be updated until the data are finalized in approximately June of the year following the year they were reported to the NNDSS database. In addition, the NNDSS data that are printed in the *MMWR* weekly or *MMWR Summary of Notifiable Diseases* represent only a subset of data transmitted to the NNDSS database. This is because criteria have been developed by CSTE and/or CDC to determine which cases to print in the *MMWR* weekly and/or *MMWR Summary of Notifiable Diseases*. The selection criteria used to assess which cases to print (or conversely, which cases to suppress from printing) can be based on one or more of the following: 1) the content of the NNDSS case confirmation status variable used to designate whether the reported case is confirmed, probable, suspect, or has unknown case classification status, 2) rare disease verification procedures, and 3) the state-specific reporting status of a disease.

13. When calculating the incidence rate of a nationally notifiable disease, the denominator may need to be adjusted to exclude the population from states where the disease was not notifiable or data were unavailable.

When calculating an incidence rate of a nationally notifiable disease, it may be appropriate to exclude from the denominator, all states that could not have contributed cases in a given year because the specific disease/condition of interest in a given year was either "not notifiable" in that state (the state was given an "NN" designation in the *MMWR* weekly or *MMWR Summary of Notifiable Diseases*, for the relevant year) or data from a specified state were "unavailable" (the state was given a designation of "NA" in the *MMWR* weekly or *MMWR Summary of Notifiable Diseases*, for the given year). In both of these situations, (designations of "NN" or "NA"), the NNDSS surveillance database does not contain case count data for the specified disease for that state and year. If the denominator of the incidence rate calculation is not adjusted in this manner, the resulting rate may be an underestimate of the incidence. All other factors that contribute to the data being an underestimate (such as under-recognition and under-reporting of disease) will also still apply. When the proposed NNDSS-based analysis is expected to span numerous years, this type of adjustment to the denominator becomes a very labor intensive effort, but may be important.

14. NNDSS data pertaining to race and ethnicity may be incomplete and may have been collected differently across the various notifiable diseases and across time.

Public health surveillance data are published for selected racial and ethnic population groups because these variables may be risk markers for certain notifiable diseases. Risk markers can identify potential risk factors for investigation in future studies. Data regarding race and ethnicity also can be useful for identifying groups to target for prevention efforts. However, caution must be used when drawing conclusions from reported data relating to race and ethnicity. Among certain races and ethnicities, there are likely to be differential patterns of access to health care, interest in seeking health care, and detection of disease that would lead to data that are not representative of disease incidence in these populations.

Other limitations regarding race and ethnicity are related to how these data are collected by surveillance programs. Race and ethnicity data are not collected uniformly for all diseases. Race and ethnicity data are collected using a single race/ethnicity variable for AIDS/HIV, sexually transmitted diseases (STD), and tuberculosis (TB). For nationally notifiable diseases other than AIDS/HIV, STD, and TB, race and ethnicity were collected as a single variable before 1992, using a different coding scheme than used for AIDS/HIV, STD, and TB. Then, beginning in 1992, race and ethnicity data for nationally notifiable diseases other than AIDS/HIV, STD, and TB were collected as two separate variables. Data on race and ethnicity are difficult to compare when the data are collected in different formats either across diseases or across time periods. When doing a data analysis spanning across 1992 for nationally notifiable diseases other than AIDS/HIV, STD, and TB, one will need to convert race and ethnicity categories to one format, which may tend to introduce more unknown values for race and ethnicity before 1992 than in 1992 and thereafter. In addition, although standard procedures exist for collecting and classifying race and ethnicity data, these procedures may not always be followed, potentially impacting the quality and consistency of the data collected. For example, while it is recommended that race and ethnicity be self-reported by the case-patient, it is possible that some data on race and ethnicity are inferred from the surname of the individual or through other means, such as by designation of the health care provider. The accuracy and validity of race and ethnicity data collected via these different methods may vary substantially, thereby impacting the quality of the data collected. Lastly, there is variation between states in the completeness of race and ethnicity reporting.

15. States may batch report cases to the NNDSS database which may impact analyses of timeliness. In addition, some administrative procedures used in processing the NNDSS data may impact analyses of timeliness.

Because of surveillance staff work schedules and reporting priorities, states may batch report surveillance data to CDC at various times of the year and/or at the year's end. Batched reporting may limit the utility of the surveillance database for analyses involving time trends, seasonality, and calculations of reporting delays.

The re-use of case identifiers by states may also limit the utility of the NNDSS data for timeliness analyses. The first report of a case to CDC generates and associates a CDC date stamp with the record. A subsequent report of the same case identifier with a different person will cause the surveillance data from the first record to be updated with

the newly reported information, but the initial CDC date stamp is not overwritten. Therefore, it is possible for the initial CDC date stamp to reflect an earlier date than the event date (onset date, date of diagnosis, etc) associated with the new case. This error is introduced as an artifact of state and CDC administrative procedures. For this reason, CDC has requested that states not re-use case identifiers.

16. The NNDSS data may undergo limited logical error checking.

CDC performs limited logical error checking at the time the NNDSS data are submitted to CDC. CDC surveillance staff will contact states when potential errors are identified and will work with states to help resolve data errors. CDC also includes error checking utilities in CDC-supplied surveillance software when data are entered in the states. Some states do not use CDC-supplied software and are responsible for developing their own error checking routines. Although CDC can help to identify data errors, it is the sole responsibility of the state to ensure the accuracy and validity of the reported data.

17. **Varicella (chickenpox) and other non-notifiable diseases of public health importance may be transmitted by states to CDC for the national surveillance of these diseases in addition to the diseases designated as nationally notifiable.**

CSTE recommends national reporting of varicella to CDC. Data on other non-notifiable diseases may be transmitted to CDC for various reasons. Reasons for this may include a desire to report cases a health care provider reported to the state that were considered important from a public health perspective even though a formalized surveillance system didn't exist for that condition and the health care provider was not mandated to report the case. Another reason for reporting could be a desire to have CDC document cases of a new or acute health condition before the disease/condition can be made nationally notifiable. In general, non-notifiable diseases data may reflect reports of conditions that are reportable in a state although not at the national level or these data may reflect reports of conditions that are considered of public health importance but not notifiable at either the state or national level. Frequently, years during which a disease is nationally notifiable are preceded by years during which states reported cases of the non-notifiable disease.

C:\Documents and Settings\ald7\Local Settings\Temporary Internet Files\OLK20\what data users should know.wpd
Last revised March 21, 2001