# Standardization of Core Surveillance Data for Transmission to CDC

## Background

### Objectives

CSTE's Data Standardization Workgroup (DSWG) partnered with CDC to support NNDSS efforts towards standardization and harmonization of public health case surveillance data classes and data elements, for uses ranging from the initial reports received by public health agencies to case notification to CDC. The standardization of core data for surveillance will help to:

- Create standards for data across the case surveillance ecosystem;
- Impact and benefit jurisdiction surveillance systems and processes, including interjurisdictional sharing and transfer of case data;
- Coordinate with the CDC's emergency response case data element project; and
- Inform the new data dictionary for case notification to CDC.

Data standardization is an important component of the Data Modernization Initiative.

The DSWG efforts focus specifically on identifying and standardizing the core data elements for surveillance *to include in transmission to CDC*.

### What are Core Data for Case Surveillance?

Core surveillance data refers to data elements that are collected and used across many conditions and are not specific to any single condition or condition family. Defining these elements in a standardized manner may facilitate interjurisdictional data exchange and interoperability and accelerate the development of data collection instruments for emerging conditions.

**The DSWG focused standardization efforts on identifying which data elements in each data class should be included in routine case notifications to CDC.** Not all core surveillance data elements are necessarily appropriate for sending to CDC, and the determination of which elements qualify for routine case notifications was central to the discussion around each data element.

For each identified core data element, the DSWG sought to find common definitions, value sets/code sets, structure, and interpretation notes. Identification of a data element as part of core surveillance data does *not* require that that element be populated for every case; depending on the condition or the circumstances of an individual case, it may be appropriate for the data element to be left blank if the information is not available, not a priority, or not applicable to that condition.

With this focus on core surveillance data for routine case notifications to CDC, the standardization of many data elements that are critical to case-based surveillance across conditions and/or

DRAFT

interjurisdictional data exchange, but that are *not* appropriate for transmission to CDC (e.g., personally identifying information), was out of scope. Furthermore, data elements relevant for only a small number of conditions and data elements that are primarily used for special studies or research purposes were excluded from consideration. Any guidelines pertaining to specific conditions were considered out of scope for the discussion, including, for example, condition-specific timeframes or value sets that may be used for core data elements. Also out of scope were the mechanisms that should be used for sending or exchanging data, guidance on when case notifications should be sent, specific implementation guidelines, and the list of data elements to be sent to CDC for emergency response.

## Data Sources and Existing Standards

As the public health community develops standard definitions and implementations of these core data elements, we want to utilize and align with other data standardization efforts within the public health and healthcare ecosystems. Public health agencies (PHAs) continue to receive manually generated reports from healthcare entities and laboratories and frequently conduct case investigations to collect detailed information on reported cases. However, electronic data submissions are an increasingly vital source of data for public health. As we explored the existing data standards available within the current healthcare and public health landscape for each data class, we considered where public health surveillance can harmonize with fields and vocabulary already defined through other efforts and where further definition or supplementation is needed for public health purposes.

One increasingly common method of receiving data for public health surveillance is through **electronic case reporting (eCR)**, in which a case report can be automatically triggered and sent from an electronic health record (EHR) to public health, using either Clinical Document Architecture (CDA) or Fast Health Interoperability Resources (FHIR) standards. When electronic initial case reports (eICRs) are triggered from EHRs, the decision support service (housed within the Reportable Condition Knowledge Management System [RCKMS]) routes the files to the appropriate PHAs based on reporting criteria operationalized using different types of value sets, such as diagnoses and problems, laboratory tests, laboratory results, organisms, and medications. The value sets consist of the various codes that might be found in the electronic case reports.

Even as PHAs increase their reliance on data coming through eCR, one consideration is that implementations vary in their conformity to the standard. Although the implementation guide has attempted to standardize the process, individual EHR vendors and healthcare organizations may deviate from the guide in such a way that the data are not sent in the expected locations or using the expected format. Throughout this document, when the eCR data source is discussed for the various data elements, the assumption is that the electronic case reports are formatted and structured as described in the accepted implementation guide.

**Electronic laboratory reporting (ELR)** is another method of receiving public health surveillance data that stems from laboratory results. Laboratories are required to notify PHAs of reportable conditions, and while these reports were historically sent via manual processes like mail or fax, ELR allows for automation of these reports. ELR messages are typically formatted using the HL7 version 2 standard, which contributes to interoperability.

Relevant elements from **message mapping guides (MMGs)** were also reviewed as part of this effort. MMGs serve as the implementation guides for PHAs when implementing HL7 case notifications to CDC.

DRAFT

The Generic version 2 (Gen V2) MMG and/or condition-specific MMGs were referenced, depending on where relevant data elements were found. Unlike the data sources and standards described above, the MMGs do not represent a data set or data standard that will help populate data *within* a jurisdiction's surveillance system. Rather, the MMG review helped to represent a) elements and formats to which PHAs can currently map or b) the condition-specific case report forms used by CDC programs. While the MMGs helped inform the workgroup discussion, they were not considered a fixed reference point, as different MMGs and elements have been implemented to varying extents, the MMGs may not be harmonized across conditions, and needs may also have changed over time.

The **United States Core Data for Interoperability (USCDI)** is a standardized set of health data classes and constituent data elements for nationwide, interoperable health information exchange that is managed by the Office of the National Coordinator for Health Information Technology (ONC). Updated versions of the standard are finalized on an annual basis, reflecting the submissions made by interested parties to ONC during an open comment period. To date, only version 1 has been included in certification criteria for electronic health records, though a new rule may advance the criteria to version 3. Submitted data elements are evaluated and classified into one of three levels based on published criteria around overall value, maturity, and known challenges to implementation. Data elements classified as Level 0 or Level 1 are still relatively primitive in the areas of standardization and electronic exchange, while Level 2 data elements are more mature in these areas and likely to be included in a subsequent version. USCDI+, an extension to the existing USCDI that will support the identification and establishment of domain- or program-specific data elements and value sets, including a set for public health, is also under development but does not yet have any specified relationship to standards development or certification criteria for EHRs.

## Process

To ensure that as many relevant stakeholders as possible were included in this core data standardization effort, the project was publicized to a wide audience, including public health agency staff, CSTE committee members and CDC program staff, and on various calls, including the monthly eSHARE webinar. Workgroup members were also encouraged to invite any colleagues that might be interested in the topics that were to be discussed in upcoming meetings.

When the partnership for standardization of core data was introduced to the workgroup, a set of data elements and data classes was presented to the workgroup to provide an idea of the types of surveillance data that may need further standardization work. Various attributes of these data classes were considered, including the relevance of the data element across multiple conditions, the importance of the data element for public health action, and the level of standardization already seen for the data element in public health agencies. Initial priority data classes were identified to start the content-based discussion.

The first two data classes that were discussed, Hospitalization and Travel, were addressed in the standard monthly DSWG meeting. However, in order to support a faster timeline for discussing and reaching consensus on standardized definitions for all of the data classes identified, a dual Data Element-Specific Team (DEST) approach was established for all subsequent data classes. These teams each convened monthly and provided a more focused environment for team members to discuss data elements and arrive at consensus definitions for the elements identified as core data for surveillance. DSWG members and any other interested parties were encouraged to sign up to participate in DESTs as

DRAFT

they were scheduled, with the understanding that team members were expected to attend the monthly meetings regularly and to participate actively during and between those meetings in the form of providing timely responses to polls conducted and thoughtful input and comments on any draft documentation that was developed by the DEST facilitators.

Prior to the start of the deliberations for each data class, background information was gathered and presented to the workgroup members for review and level-setting. This information consisted of the existing data elements in the data class that are currently included in NNDSS MMGs, as well as details on how those elements are defined or represented in any applicable national standards, such as eCR, ELR, and USCDI.

Throughout the discussions, workgroup members were polled during and between the monthly workgroup calls to determine the extent of consensus around various topics related to the data element being discussed. The goal was to reach 80% or more agreement, while identifying and discussing any major points of dissention. Questions that received at least 80% agreement were considered resolved with "strong support". If less agreement was indicated, then additional discussion followed to work through the sticking points towards resolution. For some questions, the target agreement of 80% could not be reached even after additional discussion and/or further clarification of the question, and the majority decision was taken as having "weak support".

The chapters that follow for each data class include a table of the data elements under that data class recommended by the DSWG to be included as core data for surveillance in routine case notifications to CDC (named Table x-1). The table includes the names and definitions of each data element, along with the columns below defined for each data element:

- CDC Priority: Indicates the level of requirement specified by CDC for the data element, in alignment with the prioritization format used in the MMGs. Priority values include: R = Required, 1 = highest priority for reporting, 2 = high priority, and 3 = lower priority for reporting.
- Standard Code: The recommended standard code for mapping the data element.
- Value Set or Coding System: The set of allowed values or the coding system to which the allowed values belong.
- Abstracted / Derived / Created: Captures if a data element can be directly abstracted from a data source, is derived by the jurisdiction based on other data elements collected from the data source, or is created locally within the jurisdiction (e.g., unique case ID).
- Recommendations for Input Data Source(s): Expected source(s) of information for the data element to public health.

DRAFT

# Appendix A

**List of Abbreviations**

| Abbreviation | Definition |
|---|---|
| CDA | Clinical Document Architecture |
| CDC | Centers for Disease Control and Prevention |
| CSTE | Council of State and Territorial Epidemiologists |
| DEST | Data Element-Specific Team |
| DMI | Data Modernization Initiative |
| DSWG | Data Standardization Workgroup |
| eCR | Electronic Case Reporting |
| EHR | Electronic Health Record |
| eICR | Electronic Initial Case Report |
| ELR | Electronic Laboratory Reporting |
| FHIR | Fast Health Interoperability Resources |
| MMG | Message Mapping Guide |
| NNDSS | National Notifiable Diseases Surveillance System |
| ONC | Office of the National Coordinator for Health Information Technology |
| PHA | Public Health Agency |
| RCKMS | Reportable Condition Knowledge Management System |
| USCDI | United States Core Data for Interoperability |

DRAFT

# Appendix B

**Potential Data Elements and Data Classes for Core Data Standardization**

- Case Identifiers, including CDC Emergency Response Case ID
- Disability
- Health Equity/Social Determinants of Health (SDOH)
- Hospitalization
- Immunization History
- Industry and Occupation
- Lab Identifiers (to link case and lab data)
- Laboratory Observations
- Medications
- Race/Ethnicity
- Reporting Source and Information Source
- Pregnancy
- Sexual Orientation and Gender Identity (SOGI)
- Signs and Symptoms/Clinical Manifestations
- Travel
- Tribal Affiliation
- Underlying Conditions, Exposures, and Risk Factors

DRAFT