



Core Case Notification Data Element Recommendations: Signs, Symptoms, and Clinical Manifestations

This document was developed by the Data Standardization Work Group (DSWG) of the Council of State and Territorial Epidemiologists (CSTE) Surveillance Practice and Implementation Subcommittee (SPIS) as part of a larger data harmonization effort. It will be published and maintained in a repository on the CSTE website.

The DSWG identified and standardized the data elements to utilize in the transmission of core surveillance case data to the Centers for Disease Control and Prevention (CDC). Core surveillance data classes and data elements are those that apply across many conditions and, in general, are not specific to any single condition or condition family. In each recommendation document, the DSWG provides common definitions, value sets/code sets, structure, and interpretation notes for each identified data element that is designated as core. The DSWG does not provide questions which would be used to populate the recommended data elements. More information about the DSWG and data harmonization can be found [here](#).

The structure of this document is as follows:

- [Background](#)
- [Core Surveillance Signs, Symptoms, and Clinical Manifestations Data Elements for Routine Case Notifications](#)
- [Appendix A. Input Data Standards and Data Sources for Signs, Symptoms, and Clinical Manifestations Data](#)
- [Appendix B. Data Class Structure in Message Mapping Guides](#)

Background

Signs, symptoms, and clinical manifestations are often overlapping concepts that are key to public health surveillance to describe cases of disease, establish trends, and identify potential outbreaks. Signs and symptoms are conceptually distinct in that signs typically refer to objective evidence of disease, especially as observed by the physician, whereas symptoms are reported subjectively by the patient.¹ Clinical manifestations are more all-encompassing and may be used interchangeably for either term.

It is essential to support the collection of as much information as possible during a case investigation. While questions about a person's clinical manifestations may be listed out individually in a surveillance system, the Message Mapping Guides (MMGs) for CDC case notifications most often utilize a repeating group structure for either "signs and symptoms" or "clinical manifestations". The repeating group structure is essentially the same, regardless of whether it is called signs and symptoms (Figure 1) or clinical manifestations (Figure 2). There are no MMGs that include both types of repeating groups.

The primary data element in the repeating group has a value set that captures the signs and symptoms or clinical manifestations relevant to the condition being reported. In either repeating group, the concepts included in the value sets range from more common (e.g., fever, headache, cough) to more specific (e.g., thrombocytopenic disorder, submandibular salivary gland swelling).

¹ Merriam-Webster.com. 2024. <https://www.merriam-webster.com> (16 July 2024)



Figure 1. Data elements in Signs and Symptoms Repeating Group from the Mumps MMG.

Data Element (DE) Name	DE Identifier Sent in HL7 Message	Data Element Description	Value Set Name (VADS Hyperlink)
START: Signs and Symptoms Repeating Group			
Signs and Symptoms	56831-1	Sign and symptoms associated with the illness being reported	Signs and Symptoms (Mumps)
Signs and Symptoms Indicator	INV919	Indicator for associated signs and symptoms	Yes No Unknown (YNU)
END: Signs and Symptoms Repeating Group			

Figure 2. Data elements in the Clinical Manifestations Repeating Group from the Lyme Disease MMG.

Data Element (DE) Name	DE Identifier Sent in HL7 Message	Data Element Description	Value Set Name (VADS Hyperlink)
START: Clinical Manifestations Repeating Group			
Clinical Manifestation	INV929	Clinical manifestation of Lyme disease	Clinical Manifestations (Lyme)
Clinical Manifestation Indicator	INV930	For each clinical manifestation reported, indicate whether the subject developed the specified manifestation as a result of the illness.	Yes No Unknown (YNU)
END: Clinical Manifestations Repeating Group			

The use of individual data elements for any type of clinical manifestation is an alternative structure to the repeating group. This approach is used in the Arboviral MMG, in which a “Clinical Section” contains approximately 30 individual questions about various clinical findings, built as a series of discrete data elements. Each data element is associated with a Yes/No/Unknown value set to indicate whether that clinical manifestation was present (Figure 3).

Figure 3. Sample of the data elements in the Clinical Section of the Arboviral MMG.

PHIN Variable	Data Element (DE) Name	Data Element Description	Value Set Name (VADS Hyperlink)
START: Clinical Section			
386661006	Fever	Fever	Yes No Unknown (YNU)
43724002	Chills or Rigors	Chills/ Rigors	Yes No Unknown (YNU)
271795006	Fatigue or Malaise	Fatigue and Malaise	Yes No Unknown (YNU)
271807003	Rash	Rash	Yes No Unknown (YNU)
68962001	Myalgia	Myalgia	Yes No Unknown (YNU)
57676002	Arthralgia	Arthralgia	Yes No Unknown (YNU)
3723001	Arthritis	Arthritis	Yes No Unknown (YNU)
25064002	Headache	Headache	Yes No Unknown (YNU)
44695005	Paralysis or Paresis	Paralysis or Paresis	Yes No Unknown (YNU)
161882006	Stiff Neck	Stiff Neck	Yes No Unknown (YNU)



A repeating element structure is a third approach—one not commonly used for signs, symptoms, or clinical manifestations. Repeating elements use a single data element “question” with a value set that lists all clinical manifestations appropriate to the condition, where the user would select all that apply to indicate a positive response. An example is a data element question of “Did the person have any of the following clinical manifestations?”, with a value set including “Fever,” “Cough,” and “Sore throat.” A more comprehensive description of these structures can be found in Appendix B of this document.

To support further standardization of the transmission of signs, symptoms, and clinical manifestations, the DSWG considered the following questions:

- Is it necessary to maintain signs and symptoms and clinical manifestations as distinct repeating groups? If not, how should a single data class representing this concept be named and defined?
- For this data class, what is the preferred structure?
- Is there a need for an overall data element representing whether or not the patient experienced any clinical manifestations?
- Should information such as condition complications, sequelae, and infection types be defined and standardized as additional core data elements, or should they be implemented in a condition-specific manner as needed?



Core Surveillance Signs, Symptoms, and Clinical Manifestations Data Elements for Routine Case Notifications

The DSWG recommended that the data class and data elements below be defined as core surveillance data for case notifications to the CDC. For each, there is no pre-defined default value for what should be transmitted when the data element has not been populated. Refer to Tables 1-2 below for definitions and additional information for each data element.

- [Signs and Symptoms](#) (data class)
- [Signs and Symptoms Data Class Indicator](#)*

Data elements marked with an asterisk (*) did not reach 80% consensus within the group but were supported by more than half of the participants. These are included here as core data elements with weak support from the DSWG.

The DSWG agreed with consensus that Signs and Symptoms should be a core data class to ensure that there is a standard place in message mapping guides for sending information on signs, symptoms, and other clinical manifestations for all applicable conditions. The workgroup agreed that while a simplified data class name, "Clinical Manifestations," would be conceptually inclusive of signs and symptoms, as well as other manifestations (such as a finding like "high white blood cell count"), there is a benefit to maintaining use of the name "Signs and Symptoms" for this data class as it is more colloquial and will lead to smoother adoption. Clinical manifestations of interest for a condition vary, and the relevant clinical manifestations to be captured for each condition will be defined by CDC programs.

The structure of the Signs and Symptoms data class was discussed in conjunction with the data classes of Exposure, Risks, and Underlying Conditions; the outcomes of that call are summarized in Appendix B. Among the outcomes outlined in [Appendix B](#), participants from the CDC expressed a strong preference for the repeating group structure, as updates can more simply be made to a value set. Among participants from jurisdictions, preference was split between a structure based on a repeating group, repeating elements, and individual questions, likely reflecting differences in surveillance systems used across the US. Additional concern was raised that, while a repeating group structure has its advantages, it also has the potential to expand indefinitely, creating new coding complications. Ultimately, the workgroup agreed that consistency in structure across conditions for a data class is most important, more than what the exact structure is. A majority of the group (74%) supported transmitting "No" and "Unknown" responses for each sign, symptom, and clinical manifestation due to the level of detail it provides. Given the benefits outlined above for the flexibility of the repeating group format and ability to transmit negative or unknown responses, this is likely the most advantageous structure for this data class (see Table 1).

Separate from the information captured under the umbrella of Signs and Symptoms, the workgroup also discussed information currently transmitted under the categories of Complications, Clinical Findings, and Infection Type. These data classes convey information similar to Signs and Symptoms but routinely captured as separate data elements within MMGs. The workgroup concluded that these data classes should be handled in a condition-specific manner, and the structure of these data elements standardized across MMGs when used. Several other signs and symptoms-related data elements are included in existing MMGs but are only relevant to specific conditions and would be outside the scope of core surveillance data for case notifications to CDC. Such data elements were mentioned briefly during workgroup meetings but not discussed at length in the context of core surveillance data.

Table 1. Core Surveillance Data Class for Routine CDC Case Notification – Signs and Symptoms

Signs and Symptoms	
Definition	Signs, symptoms, and other clinical manifestations associated with the condition being reported
Standard Code for Transmission to CDC	Utilizing the repeating group format, Signs and Symptoms would be transmitted as two data elements: - Signs and Symptoms- LOINC: 56831-1 Problem-associated signs and symptoms - Signs and Symptoms Indicator- CDC Local Code: INV919 Indicator for associated signs and symptoms See Interpretation Notes below.
Value Set or Coding System for Transmission to CDC	Individual signs, symptoms, and clinical manifestations should be represented using SNOMED-CT when available. Local codes may be created as needed. These value sets will be defined by CDC programs. The Signs and Symptoms Indicator would take values of Yes/No/Unknown.
Abstracted/ Derived/ Created	Abstracted
Derived from (examples)	N/A
Recommendations for Input Data Source(s)	Case investigation eCR
Notes	This data class includes transmission of signs, symptoms, and any other clinical manifestations associated with a condition. The value set for which signs, symptoms, and other clinical manifestations are associated with a condition is condition dependent. This value set may contain a response option for “other, please specify” to capture any associated signs, symptoms, or other clinical manifestations.
Data Source Details*	
eCR CDA	HL7 CDA R2 Implementation Guide: Public Health Case Report, Release 2 STU Release 1.1 - US Realm - ClinicalDocument/component/Problem Section/Problem Concern Act/Problem Observation/code - ClinicalDocument/component/Problem Section/Encounter Diagnosis/Problem Observation/code HL7 CDA R2 Implementation Guide: Public Health Case Report – eICR Release 2, STU Release 3.1 - US Realm Available as a coded field: - ClinicalDocument/component/Problem Section/Problem Concern Act/Problem Observation/code - ClinicalDocument/component/Problem Section/Encounter Diagnosis/Problem Observation/code Uncoded text that may contain relevant information: - ClinicalDocument/component/Reason for Visit Section/text - ClinicalDocument/component/Review of Systems Section/text
eCR FHIR	Resource Profile: eICR Composition - Composition.section:sliceProblemSection - Composition.section:sliceHistoryOfPresentIllnessSection - Composition.section:sliceReviewOfSystemsSection Resource Profile: US Public Health Condition - Condition.code
ELR	N/A
USCDI v6	Problems

	• Definition: Condition, diagnosis, or reason for seeking medical attention. • Applicable Vocabulary Standard(s): SNOMED-CT U.S. Edition March 2025 Release, ICD-10-CM 2025 • Part of the Problems data class.
USCDI+ Public Health: Case Reporting Use Case	Symptoms (list) • Definition: List of patient symptoms (structured) for the reportable event. • Part of the Problems data class.

* Additional information on input data sources for the Signs and Symptoms data class can be found in Appendix A.

Interpretation Notes – Clinical Manifestations Data Class

An example of the two data element repeating group where responses of “Yes” for fever and “No” for cough are transmitted:

```
OBX|nn|CWE|56831-1^Signs and Symptoms^LN|1|386661006^Fever^SCT|||||F
OBX|nn|CWE|INV919^Signs and Symptoms Indicator^PHINQUESTION|1|Y^Yes^HL70136|||||F
OBX|nn|CWE|56831-1^Signs and Symptoms^LN|2|49727002^Cough^SCT|||||F
OBX|nn|CWE|INV919^Signs and Symptoms Indicator^PHINQUESTION|2|N^No^HL70136|||||F
```

Table 2. Core Surveillance Data Element for Routine CDC Case Notification – Signs and Symptoms Data Class Indicator

Signs and Symptoms Data Class Indicator	
Definition	Indicator that the person experienced any listed signs, symptoms, or other clinical manifestations associated with the condition being reported
Standard Code for Transmission to CDC	TBD
Value Set or Coding System for Transmission to CDC	Yes/No/Unknown
Abstracted/ Derived/ Created	Derived
Derived from (examples)	If a person has any listed signs, symptoms, or other clinical manifestations associated with the condition being reported, the response would be “Yes”. If the person does not have any of the listed signs, symptoms, or other clinical manifestations, this response is “No”. If this is not known, the response is “Unknown”.
Recommendations for Input Data Source(s)	Case Investigation eCR
Notes	This data element is a summative indicator that communicates if the person has any signs, symptoms, or other clinical manifestations associated with the condition as listed in the Signs and Symptoms value set. If the Signs and Symptoms value set includes a response option for collecting “other” manifestations, the definition of this data element would encompass any signs, symptoms, or other clinical manifestations associated with the condition regardless of explicit inclusion in the value set.
Data Source Details	
eCR CDA	N/A
eCR FHIR	N/A
ELR	N/A
USCDI v6	N/A
USCDI+ Public Health: Case Reporting Use Case	N/A

Appendix A. Input Data Standards and Data Sources for Signs, Symptoms, and Clinical Manifestations Data

Electronic Case Reporting (eCR)

Data related to signs, symptoms, or other clinical manifestations of a condition may be captured in several locations within an eCR transmission. The most common templates used to transmit this data in an electronic initial case report (eICR) are as follows.

- The Problem Observation template reflects a discrete observation about a patient's problem using SNOMED or ICD10 codes. This template is used across multiple sections within the eICR, and the ones most relevant to the current condition are the Problem Concern Act and the Encounter Diagnosis.
- Codes in the Problem Concern Act may persist within eICR documents as active, even if entered during previous encounters. Utilizing the fields of effectiveTime/high and effectiveTime/low may be necessary to ensure listed codes are relevant to the current instance of an illness.
- Within the STU Release 3.1 additional document sections may transmit text-based narratives that contain information about signs, symptoms, or other clinical manifestations. These include the Reason for Visit Section and Review of Systems Section.

Electronic Laboratory Reporting (ELR)

Information on signs, symptoms, or other clinical manifestations is typically not collected or transmitted in ELR. Relevant content may be found in or derived from the Reason for Study field (OBR-31), but this field is unlikely to be comprehensive and may contain information unrelated to signs, symptoms, or other clinical manifestations (such as exposures or diagnoses of medical conditions).

United States Core Data for Interoperability (USCDI) and USCDI+

The Problems data class contains the following data elements that may relate to the transmission of signs, symptoms, and other clinical manifestations information:

Table 3. Data elements within the Problems data class of USCDI or USCDI+ that relate to the transmission of signs, symptoms, and other clinical manifestations.

Data Element	Definition	Status	USCDI+
Problems	Condition, diagnosis, or reason for seeking medical attention.	USCDI V1 and subsequent	No
Date of Diagnosis	Date of first determination by a qualified professional of the presence of a problem or condition affecting a patient.	USCDI v2 and subsequent	Yes – Case Reporting
Date of Resolution	Date of subsiding or termination of a symptom, problem, or condition.	USCDI v2 and subsequent	No
Date of Onset	Date or estimated date when signs or symptoms of a condition began. Usage note: This may be a specific day, week, month, or year, or it may be an estimate.	USCDI v6	Yes – Case Reporting
Symptoms (list)	List of patient symptoms (structured) for the reportable event.	N/A	Yes – Case Reporting
Date of First Noted/Documented	Date of first noted/documentated by a qualified professional of the presence of a problem condition affecting a patient.	Level 0	No

The USCDI guidance indicates that the applicable value sets for Problems should utilize the SNOMED-CT and ICD-10-CM coding systems. Although Date of Diagnosis, Date of Onset, and Date of Resolution

are not recommended as core data elements here, they may contain information that can help interpret if a problem is historical or associated with a current condition.

In USCDI+, the Problems data class has a data element “Symptoms (list)” that replaces the data element “Problems”. This data element is defined as a structured list of patient symptoms for the reportable event. Date of Onset is also included in USCDI+ for Case Reporting, where it is defined as the date when a problem/condition started.

Appendix B. Data Class Structure in Message Mapping Guides

Many data elements have one-to-one relationships with values. Other data elements have one-to-several relationships. This appendix will define the data transmission structures associated with the different relationships and outline how these relate to the Exposure, Risk Factor, and Underlying Health Conditions (ERU) and Signs, Symptoms, and Clinical Manifestations (SSCM) data classes.

NOTE: Current message mapping guides (MMGs) are written to align with HL7 message structure, and the following descriptions will use this format as the basis for examples. However, these structures can be translated to other messaging formats if necessary.

General Outline of Data Element Structure

- **Individual elements** have one specific piece of information paired with one value set. For example, a person only has one country of birth. This information is transmitted with a data element code paired with the value. In this example, Country of Birth (78746-5) and Mexico (MEX) are paired to transmit the information that the person was born in Mexico.
- **Repeating elements** allow for one instance of a data element code to be associated with more than one value. An example of this type of data element in NNDSS is Race Category. The Race Category (PID-10) is sent one time in an HL7 message with greater than or equal to one value. For example, for a person who identifies as both African American and Asian, two codes can be sent in the HL7 PID-10 segment using a delimiter (~) to distinguish the two codes as shown here: 2054-5^Black or African American~2028-9^Asian.
In terms of data collection, repeating elements align neatly with a select-all-that-apply format. Using the Race Category example, a form might list all Race Categories each with a checkbox. A person completing the form ticks the relevant boxes. A key limitation of this structure in terms of case notifications is the lack of clarity in the unchecked box and whether it conveys “Unknown” or “No”. While this might not be a limitation in the example where a person is self-selecting a Race Category, it might be one in other data collection scenarios. For example, if a case investigator interviews a person and asks them specifically if they have diabetes mellitus, the answer is likely to be a clear yes or no. However, if the question is not asked, the answer is neither “yes” nor “no”, but instead the information is unknown.
- **Repeating groups** allow for sets of data element codes and value codes to be repeated as necessary. This structure is most commonly used when there is a group of related data elements which naturally repeat together. One example is a group of data elements related to a single laboratory test and their associated values. The entire group of data elements will be repeated for each laboratory test transmitted in the message.

All three data structures can be used to meet the needs of case reporting, however, the structures have different strengths and limitations. Because of these, there are several times in the DSWG Core Surveillance documents when one structure or another is indicated as preferred or required.

Structure of Y/N/U Data Elements: Discussion Summary

Because of the nature of the ERU and SSCM data classes, the Data Standardization Workgroup (DSWG) conducted two special sessions, similar to those of the Data Element Specific Team (DEST) sessions, to discuss the technical implications of the data element structures. These discussions were promoted across CSTE and had a target audience of jurisdictional partners who are knowledgeable about case notification message creation, including those who might not be members of the DSWG or CSTE.

In general, participants consisted of informaticists and information technologists with knowledge of current MMGs which contain all three types of data structures outlined above: individual data elements, repeating elements, or repeating groups (also referred to as repeating blocks). These conversations specifically

focused on the ERU and SSCM data classes, but the information learned in them can be applied to most data elements which could be sent in more than one data structure.

For the purposes of this summary, a set of underlying conditions (diabetes mellitus, COPD, bipolar disorder) is used as an example to describe how ERU information might be sent in each of the three structures, as outlined below:

- **Individual elements** are built as a series of discrete data elements, each with a Yes/No/Unknown (Y/N/U) value choice. An ERU data class built with individual data elements will have several elements, each addressing a specific exposure, risk factor, or underlying condition separately. Table 7 shows how these data elements are discrete data elements with single answers.

Table 4. Underlying Condition example as individual elements

Data Element	Value Set
Diabetes mellitus	Y/N/U
COPD	Y/N/U
Bipolar disorder	Y/N/U

- **Repeating elements** use the data class label as a broad data element paired with a select-all-that-apply value set. An example is a data element question of “Underlying Health Conditions”, with a multi-select value set including “diabetes mellitus”, “COPD”, and “bipolar disorder”. Table 8 shows how these would be set up as a single select-all data element.

Table 5. Underlying Condition example as repeating elements

Data Element	Value Set
Underlying Health Conditions (select all which apply)	Diabetes mellitus COPD Bipolar disorder

- **Repeating groups:** For data classes such as ERU and SSCM, each repeating group would have a parent data element to denote the Underlying Health Condition value, for example “COPD”, and a corresponding child Yes/No/Unknown (Y/N/U) indicator to denote the presence or absence of that specific condition. This structure repeats as needed for all underlying conditions of interest for a particular condition which are listed out in the underlying condition value set. Table 9 shows how these paired data elements would repeat, with shading to highlight each paired set.

Table 6. Underlying Condition example as a repeating group of elements

Repeating Group Number	Data Element	Value Set
1	Underlying Health Conditions	Diabetes mellitus
1	Underlying Health Conditions Indicator	Y/N/U
2	Underlying Health Conditions	COPD
2	Underlying Health Conditions Indicator	Y/N/U
3	Underlying Health Conditions	Bipolar disorder
3	Underlying Health Conditions Indicator	Y/N/U

Participants from CDC strongly preferred a repeating group structure, as modifications to value sets do not require the lengthy review process that new data elements do. In the example above, this would mean that an additional underlying condition, such as Asthma, could be added to the value without extensive review. This enables rapid MMG adaptation to new developments in condition characteristics.

In addition, value set changes can generally be incorporated into MMGs in a way that prevents jurisdictions from having to re-onboard the MMG in the event of such changes.

Jurisdictional participants' structure preferences were split between repeating group, repeating element, and individual data elements. This split likely reflects differences in surveillance systems used across the US; however, it was agreed that consistency in structure across conditions for each data class is more important than prescribing the exact structure.

One potential benefit of the repeating group structure noted by CDC is that they could add additional "child" data elements to a data class relatively easily. However, jurisdictional group participants noted that a repeating group dependency structure of parent/child/.../child is more easily implemented than a parent/child/grandchild structure. Figure 4 demonstrates this with an example: a parent element in a repeating group is to select an underlying health condition (e.g., bipolar disorder), and the child element indicates if the person has a diagnosis of bipolar disorder, does not, or if it is unknown. Adding a "date of diagnosis" variable is simpler if it is not dependent on the answer of "yes" to the first question (making it a grandchild) and is instead tied directly to "bipolar disorder" (making it a second child).

Figure 4. Illustration of parent-child relationship vs. parent-child-grandchild relationship.

