Core Case Notification Data Element Recommendations: Exposures, Risk Factors, and Underlying Health Conditions

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 Exposures, Risk Factors, and Underlying Health Conditions Data Elements for Routine Case Notifications
- Appendix A. Input Data Standards and Data Sources for Immunization Data
- Appendix B. Data Class Structure in Message Mapping Guides

Background

Public health data collection often includes information on exposures, risk factors, and underlying health conditions. Exposures are generally understood to be specific or potential settings, events, or activities in which a person may have been exposed to a condition including dates and locations of exposure. Risk factors and underlying conditions broadly encompass any aspects of a person's life or environment that may place that person at higher risk of acquiring a condition or developing a more severe illness. More specifically, risk factors typically refer to behavioral, social, or situational aspects of a person's life, while underlying conditions are typically understood to be specific to a person's medical or health-related conditions. While each concept can be defined separately, as data classes, the boundaries between each class can be easily blurred.

Categorization of data elements into the above data classes in Message Mapping Guides (MMGs) is dependent on condition. It may be appropriate for the same value to be considered an exposure in one condition, a risk factor in another, and perhaps even as an underlying condition in a third. A standardized definition for each concept would facilitate the application of the appropriate categorization to each data class. For example, the following elements are included in MMGs as different data classes for different conditions:

- Residence in a long-term care facility is captured as a risk factor in the Tuberculosis (TB) MMG but as an exposure in the contact section of the Hepatitis B, acute MMG.
- Diabetes is captured as an underlying condition in the Respiratory and Invasive Bacterial Disease (RIBD) MMG, but as a risk factor in the COVID-19 MMG.

Originally released September 23, 2025

The message structure used to capture these data classes also varies by condition within current MMGs. One message structure is the repeating groups which were developed and incorporated into several MMGs to standardize the format of information transmission to CDC. An example of an Underlying Conditions (or "Underlying Causes") Repeating group can be seen in the Carbon Monoxide MMGs shown below in Figure 1. This repeating group structure allows for the same data element to be used in multiple MMGs but with different value sets – i.e., different underlying conditions in this example – relevant to the specific condition in the MMG.

Figure 1. Underlying Conditions Repeating Group data elements from the Carbon Monoxide (CO) MMG

Data Element (DE) Name	DE Identifier Sent in HL7 Message	Data Element Description	Data Type	Value Set Name (VADS Hyperlink)
START: Underlying Conditions Repeating Group				
Underlying Condition(s)	INV236	Select the patient's preexisting condition(s).	Coded	Underlying Conditions (CO)
Underlying Conditions Indicator	INV662	Indicator for underlying condition(s)	Coded	Yes No Unknown (YNU)
END: Underlying Conditions Repeating Group				

On the other hand, some MMGs use individual data elements to capture these types of data. For example, the Babesiosis MMG includes the elements in Figure 2 to collect various underlying conditions that are relevant to a case of babesiosis. Individual data elements are helpful in capturing additional information beyond a Yes/No/Unknown response, such as date or cause, either as child data elements or as additional individual data elements.

Figure 2. Individual data elements related to underlying conditions from the Babesiosis MMG.

Data Element (DE) Name	DE Identifier Sent in HL7 Message	Data Element Description	Data Type	Value Set Name (VADS Hyperlink)
Asplenic	300564004	Is the subject asplenic?	Coded	Yes No Unknown (YNU)
Splenectomy Date	82760-0	Date of subject's splenectomy	Date	N/A
Splenectomy Reason	82759-2	Why was the subject's spleen removed?	Text	N/A
History of Previous Illness	161413004	Did the subject have a history of babesiosis?	Coded	Yes No Unknown (YNU)
Patient Immunocompromised	370388006	At the time of diagnosis, was the subject immunocompromised?	Coded	Yes No Unknown (YNU)
Immunocompromised Associated Condition or Treatment	INV933	If the subject was immunocompromised, what was the associated condition or treatment?	Text	N/A

In some cases, such as risk factors in the TB MMG shown below, a combination of a repeating group and additional individual data elements is used to collect all of the desired information.

Figure 3. Risk Factor Repeating Group and individual data elements related to risk factors from the Tuberculosis MMG

Data Element (DE) Name	DE Identifier Sent in HL7 Message	Data Element Description	Data Type	Value Set Name (VADS Hyperlink)
START: RVCT & TBLISS Risk Factor Repeating Group				
Patient Epidemiological Risk Factors	INV1117	Exposed risk factors for the patient - Please provide a response for all risk factors in the value set with an associated indicator.	Coded	Epidemiological Risk Factors (TB)
Patient Epidemiological Risk Factors Indicator	INV1118	Provide a response for each value in the patient epidemiological risk factors value set.	Coded	Yes No Unknown (YNU)
END: RVCT & TBLISS Risk Factor Repeating Group				
START: RVCT & TBLISS Additional TB Risk Factors Questions				
Type of Correctional Facility	INV1119	If patient was a Resident of Correctional Facility at Diagnostic Evaluation, indicate the type of correctional facility.	Coded	Type of Correctional Facility
Type of Long-Term Care Facility	INV1120	If patient was a Resident of Long Term Care Facility at Diagnostic Evaluation, indicate the type of long term care facility.	Coded	Type of Long Term Care Facility
Smoking Status	72166-2	What is the patient's current tobacco smoking status?	Coded	Smoking Status
Patient lived outside of US for more than 2 months	INV1121	Residence or Travel in countries other than the United States, Canada, Australia, New Zealand, or countries in northern or western Europe for >60 consecutive days at any point in the patient's lifetime.	Coded	Yes No Unknown (YNU)
END: RVCT & TBLISS Additional TB Risk Factors Questions				

The above examples highlight how the same data may be structured for collection and transmission in different ways, a variation that is seen with all three data classes (a more comprehensive description of these structures can be found in Appendix B of this document). The lack of standardization in the definitions, value sets, and format of data elements results in jurisdictional difficulty related to data collection, data transmission, and implementation of MMGs.

In discussions surrounding these concepts, the DSWG considered the following questions:

- Are these three data classes distinct? If so, how?
- Which of these data classes should be considered part of core surveillance for routine case notifications to CDC?
- What is the ideal message structure? What are the pros and cons of the repeating group structure when compared to individual data elements?

Core Surveillance Data on Exposures, Risk Factors, and Underlying Health Conditions in Routine Case Notifications

The workgroup recommends that the exposure, risk factor, and underlying health condition (ERU) data classes be considered part of core surveillance for case notifications to CDC. When polled, there was consensus that these data classes are three distinct concepts and should be defined and sent separately from one another. Table 1 lists the data concept definitions which were strongly supported by the group.

Table 1. Data Class Definitions – Exposure, Risk Factor, Underlying Health Condition

Data Class Definitions	
Exposure	Within exposure window, either the person's contact with an entity (person, animal, or substance) or presence at a location where exposure to an agent could have occurred (i.e., where did pathogen come from).
Risk Factor	Characteristics, behaviors, or social/environmental factors that increase the likelihood of developing a disease or condition.
Underlying Health Condition	Pre-existing health conditions that may worsen the course or severity of the disease or health condition.

The key distinctions between the data classes are as follows:

- Exposure includes a relevant time window, which can be different depending on the condition.
- Risk factors impact acquiring infection or developing disease.
- Underlying health conditions impact the course of infection or health condition.

While not every condition will require that data for all three data classes be sent, it was agreed that a core message format should include placeholders for each. Further details on eCR, ELR, and USCDI as input data sources on exposures, risk factors, and underlying health conditions can be found in Appendix A.

Jurisdictions will not be tasked with categorizing data elements into exposures, risk factors, or underlying health conditions on their own for transmission to CDC. Instead, CDC program areas will categorize individual data elements into the three data classes on a condition-by-condition basis to ensure continuity across jurisdictions while maintaining value set flexibility. There may be cases when similar data elements are categorized into different data classes; however, these standard definitions should increase CDC program alignment and limit these occurrences.

The structure of the Exposure, Risks, and Underlying Conditions data classes was discussed in conjunction with the Clinical Manifestations data class, and the outcomes of that call are summarized in Appendix B. The workgroup determined that for the purposes of Exposures, Risks, and Underlying conditions, it is important for states to be able to send “Yes”, “No”, and “Unknown” responses for all data elements. As suggested in Appendix B, this would align with either a repeating group structure or individual data elements for all Risks, Exposures, and Underlying Conditions being requested for any condition. Repeating groups provide the most flexibility for CDC and, thus, are the preferred option for all three data classes when possible.

In addition to defining the three data classes, the workgroup also recommends the use of summative indicators for each data class. These data elements would indicate if the person had a positive answer for any of the listed exposures, risk factors, or underlying health conditions. This type of data element may enable some jurisdictions to simplify how they build the case in addition to simplifying data analyses by providing a simple Yes/No/Unknown data element for each of the three classes. These data elements are summarized in Tables 2-4.

Originally released September 23, 2025

Table 2. Core Surveillance Data Element for Routine CDC Case Notification –Exposure

Exposure	
Definition	Within exposure window, either the person's contact with an entity (person, animal, or substance) or presence at a location where exposure to an agent could have occurred (i.e., where did pathogen come from).
Standard Code for Transmission to CDC	Utilizing the repeating group format, Exposures would be transmitted as two data elements: • Exposure - TBD • Exposure Indicator- CDC Local Code: INV1086 See Interpretation Notes below
Value Set or Coding System for Transmission to CDC	Individual exposures 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 Exposure 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	The value set for exposures of interest is condition dependent. This value set may contain a response option for “other, please specify” to capture any other exposures.
Data Source Details	
eCR CDA	See Appendix A.
eCR FHIR	See Appendix A.
ELR	N/A
USCDI v6	See Appendix A.
USCDI+ Public Health: Case Reporting Use Case	See Appendix A.

Table 3. Core Surveillance Data Element for Routine CDC Case Notification – Exposure Data Class Indicator

Exposure Data Class Indicator	
Definition	Indicator that the person experienced any listed exposure 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 exposures associated with the condition being reported, the response would be “Yes”. If the person does not have any of the listed exposures, 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 exposure associated with the condition as listed in the Exposure 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

Table 4. Core Surveillance Data Element for Routine CDC Case Notification – Risk Factor

Risk Factor	
Definition	Characteristics, behaviors, or social/environmental factors that increase the likelihood of developing a disease or condition.
Standard Code for Transmission to CDC	Utilizing the repeating group format, Risk Factors would be transmitted as two data elements: • Risk Factor – TBD • Risk Factor Indicator- CDC Local Code: INV1117 Patient Epidemiological Risk Factors Indicator See Interpretation Notes below
Value Set or Coding System for Transmission to CDC	Individual risk factors 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 Risk Factor 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	The value set for risk factors of interest is condition dependent. This value set may contain a response option for “other, please specify” to capture any other risk factors.
Data Source Details	
eCR CDA	See Appendix A.
eCR FHIR	See Appendix A.
ELR	N/A
USCDI v6	See Appendix A.
USCDI+ Public Health: Case Reporting Use Case	See Appendix A.

Table 5. Core Surveillance Data Element for Routine CDC Case Notification – Risk Factor Data Class Indicator

Risk Factor Data Class Indicator	
Definition	Indicator that the person has any listed risk factor 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 risk factors associated with the condition being reported, the response would be “Yes”. If the person does not have any of the listed risk factors, 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 risk factor associated with the condition as listed in the Risk Factor 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

Table 6. Core Surveillance Data Element for Routine CDC Case Notification – Underlying Health Conditions

Underlying Health Conditions	
Definition	Pre-existing health conditions that may worsen the course or severity of the disease or health condition.
Standard Code for Transmission to CDC	Utilizing the repeating group format, Underlying Health Conditions would be transmitted as two data elements: • Underlying Health Conditions – TBD • Underlying Health Condition Indicator - CDC Local Code: INV662 Underlying Conditions Indicator See Interpretation Notes below
Value Set or Coding System for Transmission to CDC	Individual underlying health conditions 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 Underlying Health Condition 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	The value set for underlying health conditions of interest is condition dependent. This value set may contain a response option for "other, please specify" to capture any other health conditions.
Data Source Details	
eCR CDA	See Appendix A.
eCR FHIR	See Appendix A.
ELR	N/A
USCDI v6	See Appendix A.
USCDI+ Public Health: Case Reporting Use Case	See Appendix A.

Table 7. Core Surveillance Data Element for Routine CDC Case Notification – Underlying Health Condition Data Class Indicator

Underlying Health Condition Data Class Indicator	
Definition	Indicator that the person has any listed underlying health condition 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 underlying conditions associated with the condition being reported, the response would be “Yes”. If the person does not have any of the listed underlying conditions, 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 underlying health condition associated with the condition as listed in the Underlying Health Condition 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 Sources for Exposure, Risk Factor, and Underlying Health Conditions Data

Electronic Case Reporting (eCR)

Because of the broad nature of the type of data that can be included in the ERU data concepts, such data 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-CT or ICD-10-CM codes. This template is used across multiple sections within the eICR, including Past Medical History, Problem Concern Act, and Encounter Diagnosis.

- The **Exposure/Contact Information Observation** template can be used in either the Social History or Travel History sections. It represents a person's contact with another person, animal, or substance, or the person's presence at a location where exposure to an agent could have occurred; therefore, this template might capture exposures from the environment, at a specific event or location, with an animal, or with another person.

- The **Social History Observation** template is only used within the Social History section to represent a patient's occupation(s), lifestyle, and environmental health risk factors. Data exchanged using this template can include environmental exposures, risk factors and health-related behaviors such as exercise, tobacco use and exposure, and alcohol intake.

Electronic Laboratory Reporting (ELR)

Information on ERU data is typically not collected or transmitted in ELR. Although relevant content may be found in or derived from the Reason for Study field (OBR-31), caution must be taken if using these data to populate any ERU data elements. The specific SNOMED-CT or ICD-10-CM codes used in this field would likely require manual review to determine which of the three concepts apply.

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

Table 5 below contains five data elements from the Exposure/Contact Information data class in USCDI that capture exposure-related information.

Table 5. USCDI Data Elements Relevant to Exposure-Related Information

Data Element	Definition	Status	USCDI+
Exposure/Contact Source/Target Participant	The source or target of the exposure or contact. Could be person, animal, or location.	*Level 1*	Yes
Exposure/Contact Direction	Whether the direction of exposure/contact is acquisition (person is the target and another person, animal, location, etc. is the source) or transmission (patient is the source and another person, animal, location, etc. is the target).	*Level 1*	Yes
Exposure/Contact Type	The type of exposure/contact (environmental, activity, event, location, person, animal, etc.) to an agent.	*Level 1*	Yes
Exposure/Contact Agent	The possible agent of concern in the exposure/contact.	*Level 1*	Yes
Exposure/Contact Date	The date on or period in which possible exposure/contact occurred.	*Level 1*	Yes

The USCDI guidance indicates that the applicable value sets and codes for these elements are pulled from the HL7 FHIR US Public Health Exposure Contact Information Profile. These data elements are all at a USCDI maturity level 1, indicating that they are represented by a terminology standard or standards development organization balloted technical specification or implementation guide, are capture stored or accessed in at least one production EHR or HIT module, are electronically exchanged between two production EHRs or other HIT modules using available interoperability standards, and the use cases apply to several care settings or specialties.

Unlike Exposures, there is no single data class that offers data elements useful to the Risk Factors concept, but information on specific risk factors is spread across several data classes.

Table 6. USCDI Data Elements Relevant to Risk Factor-Related Information

Data Class	Data Element	Definition	Status	USCDI+
Problems	Problems	Condition, diagnosis, or reason for seeking medical attention.	Included since Version 1	No
Problems	SDOH Problems / Health Concerns*	Social Determinants of Health (SDOH)-related health concerns, conditions, or diagnoses. Examples include but are not limited to homelessness and food insecurity.	Included since Version 2	No
Health Status Assessments	Alcohol Use	Evaluation of a patient's consumption of alcohol. Examples include but are not limited to history of alcohol use, alcohol use disorder identification test and alcohol intake assessment.	Included since Version 4	No
Health Status Assessments	Disability Status*	Assessment of a patient's physical, cognitive, or psychiatric disabilities. Examples include but are not limited to American Community Survey, Veterans RAND Health Survey, and Patient-Reported Outcomes Measurement Information System (PROMIS).	Included since Version 3	Yes – Case Reporting
Health Status Assessments	Mental/ Cognitive Status	Assessment or screening for the presence of a mental or behavioral problem. Examples include but are not limited to Confusion Assessment Method (CAM) and Patient Health Questionnaire (PHQ).	Included since Version 3	No
Health Status Assessments	Substance Use	Evaluation of a patient's reported use of drugs or other substances for non-medical purposes or in excess of a valid prescription. Examples include but are not limited to substance use disorder score,	Included since Version 4	No

		and substance use knowledge assessment.		
Health Status Assessments	Smoking Status	Assessment of a patient's smoking behaviors. Examples include but are not limited to packyears and current use.	Included since Version 3	No

* Category has been considered as a distinct data class.

The data elements listed below may also contain information pertaining to risk factors. However, being categorized as either Level 2 or Level 0, they have not yet been approved for inclusion in USCDI. Data elements marked with an asterisk (*) are currently included in the Case Reporting use case for USCDI+.

- Social History data class
    - Sexual Activity (Level 2)

- Social Determinants of Health data class
    - Congregate Living (Level 0)*
    - Refugee Status (Level 0)
    - Incarceration History (Level 0)
    - Housing Instability and Homelessness (Level 0)*

Two data elements within the Problems data class, Problems and SDOH Problems / Health Concerns (Table 6, Rows 1-2), could be applied to the Underlying Health Conditions concept. Any codes from either the SNOMED-CT or ICD-10-CM vocabulary would be applicable to these data elements. The other three data elements within the Problems data class, Date of Diagnosis, Date of Onset and Date of Resolution, should be referenced to ensure that the time period during which the condition, problem, or diagnosis is present is consistent with that of being considered underlying.

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 7. 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 8. 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 9. 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.