

Meeting: CSTE Data Standardization – Exposures, Risks, Underlying Conditions -Data Element-Specific Team Meeting

Date: April 12, 2024

Attendees: 79

CSTE (2)	PHAs (29)	CDC (43)
Becky Lampkins Gillian Haney	<u>DSWG Co-Chairs</u> Molly Crockett MA	Allison Miller Andre Dailey Andria Apostolou GDIT Angie Agunloye GDIT Brenna VanFrank Danielle Tack Doris McGinnsss Erin Miller Fidel Desir Gonza Namulanda Hailey Vest Jennifer Huang Jennita Reefhuis John Gerstle Katie Dodd Kristin Delea Lauren Korhonen Laurie Barker Lesliann Helmus Luis Hernandez GDIT Marion Rice Meeyoung Park Melissa Pagaoa Monica Patton Myrna del Mar Gonzalez GDIT Nekole Locke Nicole Lindsey Nicolette Bestul Nikolay Lipskiy Noreen Kloc Peter Boersma
JMC (3)	<u>DEST Members</u> Aaron Bieringer MN Abi Collingwood UT Anna Kocharian WI Brianna Rossi VA Cher Levenson WA Dimple Patel KY EJ Klingler NYC Elizabeth Kessler NV Erica Bye MN Hollie Sheller MO Isaiah Francis NM Kristen Merrell MT Kyra Veatch IN Lakin Mauch ND Lea Moser NV Lunden Espinosa IN Maria Jaurretche DC Mario Beltran MI Mary Vance MN Misty Johnson WI Nwanne Onumah MS Rachel Klos WI Shawna Stuck GA Sophia Cantor WA Stephanie Moberg NM Steven "Ike" Eichner TX Sue Soliva MA	
Unknown (1)		
Robbie Lojo		

	Suzanne Gibbons-Burgener WI	Peter Daly Rachelle Oseni Renee Boehmert Robert Pratt Sandy Price Sara Johnston Sara Luckhaupt Scott Grytdal Shawn Arshad GDIT sophia kazakova Sandra Roush Yuritzy Glez
--	-------------------------------	---

**SOME INDIVIDUALS DIALED IN; THE NAMES ARE UNKNOWN, AND THEY ARE LEFT OFF OF THE ATTENDANCE*

HIGH-LEVEL MEETING NOTES (REFER TO MEETING SLIDES FOR MORE DETAIL AS NEEDED):

Exposures, Risks, Underlying Conditions

DEST Opportunities for Participation

- DEST sign-ups are ongoing.
- **[Action Requested]** Please sign up for upcoming data elements on the [DEST Sign-Up here](#).
- Upcoming data elements
 - Signs/Symptoms/Clinical Manifestations (Call #2 on April 19)
 - Disability
 - Identifiers (Case/Event Related/Laboratory)
 - Medication/Treatment
 - Laboratory
 - Race/Ethnicity/Tribal Affiliation
 - Social Determinants of Health
 - Reporting Source
- DEST member expectations—members are not expected to write documents, lead meetings, or prepare slides. The goal is to keep members engaged in conversations about content during and between meetings.
- Active participation in polling, document review, discussion, and other input during and between meetings is expected.

DSWG Reminders

- SOGI recommendations draft is out to the wider DSWG for comment now through April 22

- The next general DSWG Call is on April 26 at 1 pm ET

Meeting Ground Rules

- We welcome everyone's opinions and active participation from jurisdictions, CDC programs, CSTE membership, etc.
- As we work through data elements/classes, we strive for 80/20 consensus.

In-Scope vs Out-of-Scope

- Focus on core data for surveillance and transmission to the CDC within selected data classes.
- We will work to identify standard definitions and value sets within that definition of core data.
- This is not a discussion of how to collect this data, frame questions in surveys or interviews, or implement these data standards. This is also not meant to impose limitations on how jurisdictions collect data.

Data Class Review: Exposure Risk, Underlying Conditions

Data Class Review

- As we covered in February, each of these data classes (formerly referred to as data concepts) can encompass many data elements. Therefore, each of the three data classes was decided to be considered a distinct core data class. You might remember us discussing them as buckets to capture different data elements.
- Today, we will be polling on definitions for each data class.

Proposed Definitions

- Here are the proposed definitions we will poll on in a moment; we will briefly consider what makes each one distinct before polling.
- Something to think about as we head into the polling is that these definitions are not meant to provide an academic reference point or impact how data are used or analyzed. They are specifically meant to provide some guidance for future messaging guides. The idea is to pair the definitions with recommendations about structure to enable standardization when the CDC requests these types of data.

How will data elements be sorted into three data classes?

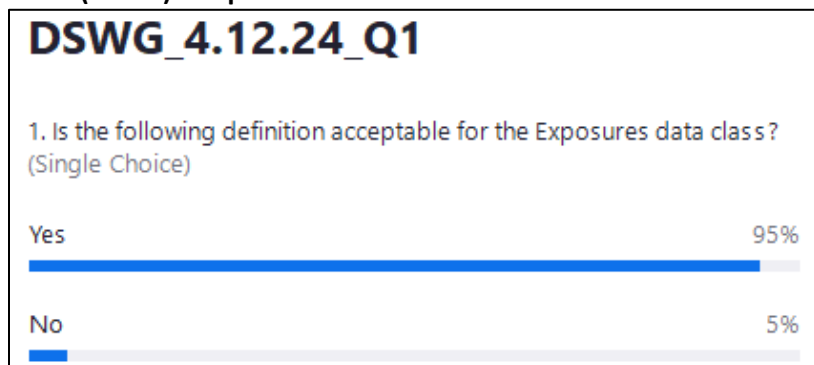
- To that end, it is important that everyone understands that within these data classes, CDC programs will oversee deciding which exposures, risks, or underlying conditions go into which data class, and the sorting may not be the same across programs (and by extension messaging guides (MMGs)).
- For example, Diabetes mellitus may continue to be considered an Underlying condition in the Respiratory and Invasive Bacterial Diseases (RIBD) MMG but could remain listed as a Risk Factor in the COVID-19 MMG.

Exposure Risk, Underlying Conditions

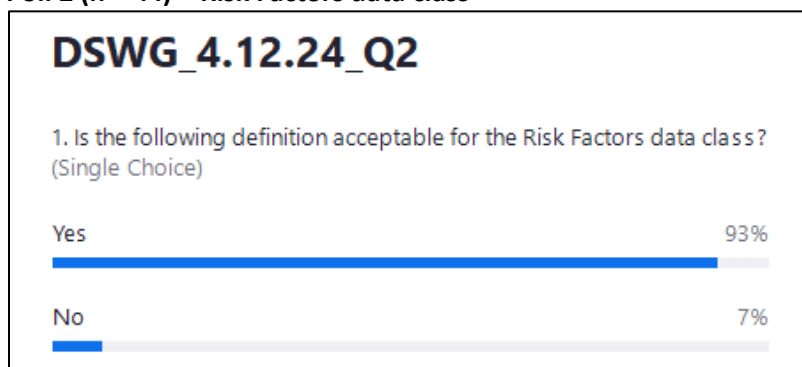
Proposed Exposure Definition

- A key point to note about this definition is that it includes a timing component that is flexible enough to allow for most if not all, condition information.

- It also incorporates objects, locations, and people, which means the exposure value set could include things such as getting a tattoo, working in a daycare, visiting a water park, etc.

Poll 1 (n = 44) – Exposures data class**Proposed Risk Factor Definition**

- Key points for this definition include that there is no timing element, allowing that to be added to the values in a value set if necessary.
- Another key point is that risk factors are related to the likelihood that someone contracts the condition, whether it is a risk factor that makes a person more vulnerable to illness or increases their likelihood of coming in contact with the condition in a way that is more general than a specific exposure might imply.

Poll 2 (n = 44) – Risk Factors data class

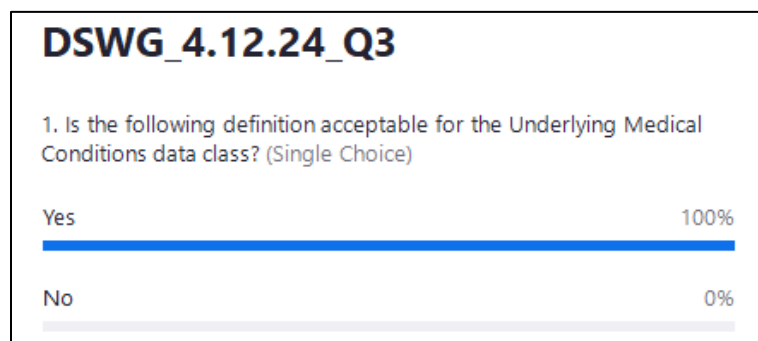
Proposed Underlying ~~Medical~~ Health Condition Definition

- Key points here are that it is asking specifically about medical conditions, and there is a reference to the medical condition impacting infection course or severity.

Discussion

- There was a verbal and chat-based discussion about how data would be categorized into each category for risk factors, exposures, or underlying conditions. JMC mentioned that the definitions are meant to provide maximum flexibility. CSTE confirmed JMC's assessment and expanded further in that these definitions cannot account for every situation. When implementation discussions begin, hopefully, the CDC or this group will be able to guide those situations for the associated value sets.
- There was additional discussion to refine the data class name and definition before voting to replace “medical” with “health. This replacement allows the definition to cover things such as mental health. The group landed on the following definition: “Pre-existing health conditions that may worsen the course or severity of the disease or health condition.”

Poll 3 (n =42) – Underlying ~~Medical~~ Health Conditions data class



ERU: Data Class Indicator

Data Class Indicator

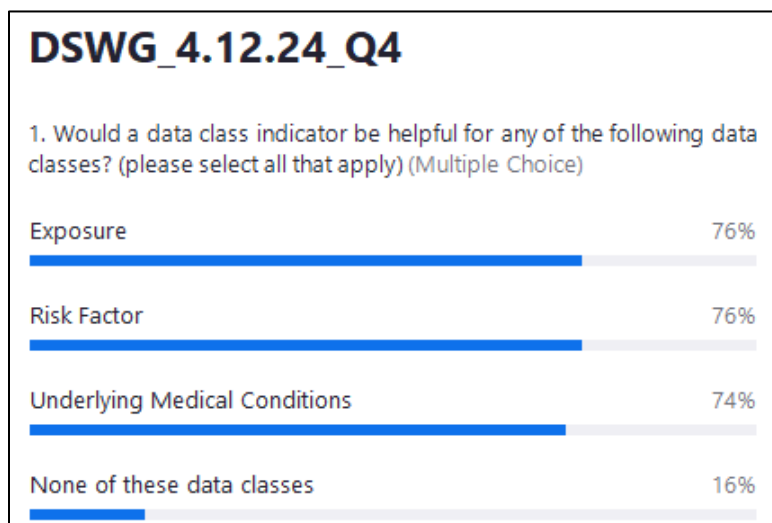
- A data class indicator is an individual data element that would capture the broader questions - “Did the subject have any exposures/risks/underlying conditions prior to illness?” with responses limited to Y/N/U.
- This was discussed during the March structure calls with the informaticists, but it was determined to be more pertinent to the epis in the group – so we are bringing it back here.
- In the example here, it is the data element included in the second row and is likely already used in some surveillance systems for these data classes or others where it might make sense.

Discussion

- Lesliann (CDC) discussed the importance of incorporating a screening element into messaging content related to data elements and informatics. This screening element would inquire about exposures, risk factors, or underlying health conditions in a general manner. If a person indicates "Yes" to exposures, risk factors, or underlying health conditions, specific content related to those will be sent. This screening element might serve as a tool to trigger relevant message segments, allowing for more targeted communication without necessarily requiring manual input from the user. This would be used as a data element to represent whether there was any data for a data class.
- Erin (JMC) clarified for the group that the informaticists wanted this to be shared back with this group to determine the specific language to use and whether this should be included or excluded as an option to have the indicator. Implementation of this would be up to jurisdiction. The current concept is that this would be a "yes", "no", or "unknown" question, which would trigger additional options for entering and sending related data.
- There were additional discussions in the chat and verbally regarding implementation, and in the interest of time, a poll was launched to gauge the group's general thoughts on having this summative data element/data class indicator.

Poll 4 (n = 38) – Would a data class indicator be helpful?

- This group's question is whether you would find this type of data class indicator helpful for any of these data classes. Please select all the answers that apply to this poll.
- If a data class clearly and strongly supports a data class indicator, it will be included in the guidance document, although how will remain to be decided.



ERU: What about Structure?

March Structure Calls- Review

- The structure we will be discussing is visualized on this slide, which we can reference later if necessary.
- The three potential structures are shown here.
- The straightforward questions are shown on the right and boxed in green. These are separate questions and are included in message mapping guides. Examples here are: Did the person enter the cattle barn at the fair? And While at the fair, did the person enter the food tent? These require quite a bit of federal approval for each question, which severely inhibits their flexibility for new MMG development and current MMG updates.
- Repeating Elements are on the top left and boxed in orange. These have a general question such as “While at the fair, did the person enter any of the following locations?” and then answers are a select all option from a value set. These limit a question to indicate a positive result with no way to send No or Unknown values.
- Repeating Blocks are at the bottom left and boxed in green. These can have the same general question, but one value set includes the locations of interest, and a paired value set allows the selection of Yes, No, or Unknown.
- These are examples of how they might look in a system and are not prescriptive.

March Structure Calls- Review (continued)

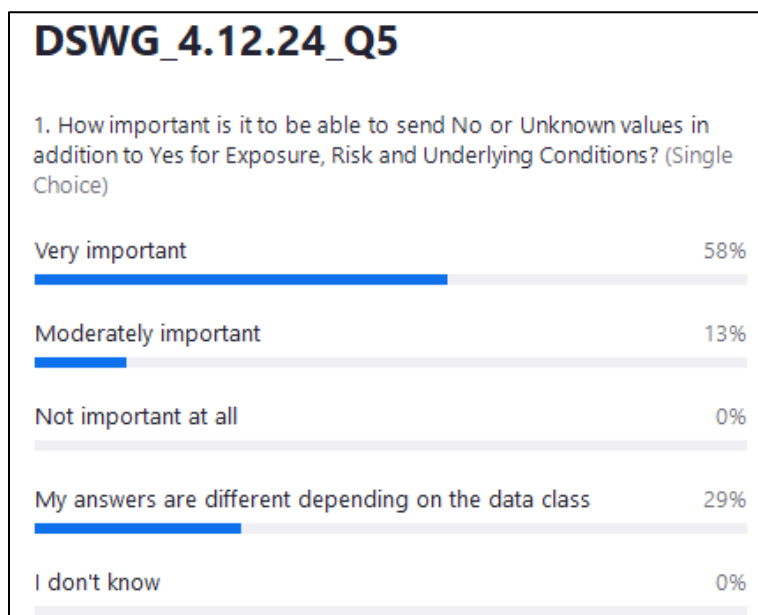
- With those options in mind, here are some, but not all, of the points made by the CDC program and jurisdictional informaticists on last month’s call.
 - It was agreed that the structure needs to be what makes sense for each data class.
 - CDC strongly preferred the repeating group due to the flexibility it allows for updating MMGs and developing new ones.
 - Informaticists from the jurisdictions were more evenly split about what works best in their system; however, they quickly pointed out that it needs to make sense for the epis and their work. However, there was concern that adding additional questions to the repeating blocks in a nested parent/child/grandchild structure would cause serious pain points for many jurisdictions.
 - Overall, there was a strong message that consistency in structure across conditions for each data class is important. This is to say that it would be beneficial to most partners if there was an exposure-repeating group for TB; for example, then there would be exposure-repeating blocks for all other conditions as well.
 - It was also acknowledged that on those calls, we are heard primarily from states with people to help with this type of work and that not all jurisdictions have equal resources and help. While Technical assistance programs are provided to states, this can only go so far as some of the work has to be done internally at the jurisdiction.

Poll 5 (n = 45) –No/unknown Values

- How important is it to be able to send “No” or “Unknown” values in addition to “Yes” for exposures, risks, and underlying conditions? Very important, moderately important, not important at all, I don’t know, OR your answer might be different depending on the data class – in which case there is an option for that.

Discussion

- Lesliann (CDC) noted that this is about the ability to send the information and not that it is a required field. If it is not relevant, it can be left blank.
- Molly (MA- Co-Chair) emphasized communication's importance in effectively conveying information. They highlighted the need for recipients to understand that leaving some options blank is acceptable, especially when dealing with uncertainties or unknowns. They mentioned challenges in filling in blanks or unknowns in certain circumstances, which could consume excessive time and effort. Molly suggested that while providing options for no and unknown responses is important, recipients shouldn't feel obligated to fill them out if it would be overly time-consuming, especially if the system does not directly support it.

**Pulse Check – Transmission/exposure questions****Discussion: Suspected point of exposure**

- Up until now, we've been referencing things that we have discussed in previous calls and having you poll on them to inform the draft documents that we have been creating and sending out for review.
- Now, we will discuss a "likely/suspected point of exposure" data element. This conversation will not lead to any core surveillance recommendations, but we will discuss it more generally.

- To start, are there any concerns or challenges related to this type of data element? These could be structural or content-related.

Discussion

- Suzanne (WI) mentioned adding a question about the likely cause of a fungal infection. WI is questioning who should determine the most likely cause: the affected individual, considering their perspective, or experts analyzing the situation comprehensively, especially in cluster investigation scenarios. She expressed difficulty in cases where diseases lack clear point source exposures. Suzanne anticipates that responses to such questions would likely be general and acknowledged the potential usefulness of the question but also suggested that individual states should bear the responsibility for interpreting epidemiological data within their jurisdiction, raising concerns about the necessity of sharing such information.
 - Misty Johnson (WI) echoed Suzanne's sentiments and discussed the process of handling disease surveillance data in Wisconsin as an informaticist. She mentioned that some forms already include a similar question, typically in a State use-only section. Local authorities collect the data, and state epidemiologists provide the final answer on where the exposure likely occurred. We could send that information if it were available in an MMG. However, if the MMG prefers a broader question about potential sources of exposure, they can still send the data as is. The State use section remains available for the exclusive use of state epidemiologists. Wisconsin's approach is flexible in incorporating such questions.
- Lesliann (CDC) questioned the frequency of reporting such information to the CDC, wondering if it's primarily for internal record-keeping or if the CDC should be informed for outbreak investigations. She also pondered alternative methods for transmitting this information to relevant authorities.
- Two individuals who work in the STI space from New York City and Philadelphia indicated in the chat that they would not send that data.
- Rachel (WI) noted that this is not about being required to send it but, moreover, having the option to send it if needed; she approved of the idea of being able to send this information if needed.
- Suzanne (WI) noted the challenges of this and considered the level of detail desired in tracking exposures. She proposed categorizing exposures based on person-to-person interactions, specific locations, or activities within those locations. She recommended starting with broader categories and refining them based on the situation. Drawing from recent experiences with Norovirus reporting, she highlighted common categories like waterborne, foodborne, and environmental exposures, suggesting further differentiation if needed. There are challenges in determining the most likely source or cause of an illness, and she questioned whether conveying such detailed information will be able to be transmitted based on the jurisdiction's comfort levels or staffing abilities.
- Lesliann (CDC) raised the question of specificity and confidence levels regarding the source of an illness. She suggested that when individuals have traveled to multiple places outside the country, it's beneficial to identify the specific location where the outbreak occurred, such as on a cruise ship. She emphasized the importance of flagging the most likely source based on a thorough investigation by local authorities, even if other exposures occurred. This approach allows for narrowing down the primary source of the illness without the need to assess confidence levels for every exposure.

- Erin Miller (CDC) was interested in gathering input from jurisdictions regarding the feasibility of updating data elements related to the likely source of exposure, as discussed by Leslieann. They raise concerns about the burden of conducting thorough investigations and then revisiting events to update the data. She invited thoughts on this matter, emphasizing the importance of considering the practicality and likelihood of jurisdictions updating such information.
 - There were several comments in the chat supporting this statement and the burden on jurisdictions for this level of data collection.

Next Steps

- JMC will now work closely with CSTE and Co-Chairs on the draft document to be sent to this team for review.
- Upcoming DESTs include
 - Signs, Symptoms, Clinical Manifestations, Call #2, on April 19
 - Disability, May 10
 - Identifiers, May 17
- The next larger DSWG call will be held on April 26