Meeting: CSTE Data Standardization – Industry and Occupation -Data Element-Specific Team Meeting
Date: November 17, 2023
Attendees: 61

CSTE (5)	PHAs (26)	CDC (23)
Becky Lampkins	Co-Chairs	Al Barskey
Cailyn Lingwall	Molly Crockett | MA	Allison Miller
Gillian Haney		Andrea Steege
Shaily Krishan	DEST Members	Antonio (Tony) Neri
Taylor Pinsent	Amanda van Rest | NV	Audrey Reichard
JMC (4)	Amy Bogucki |TN	Craig Kassinger
Erin Kough	Amy Lai | AZ	David Huang
Grace Mallard	Anna Kocharian| WI	Jenny Couse
Natalie Raketich	Cher Levenson | WA	Katie Dodd
Sam Spoto	Chris Dumas | SD	Kevin Lan
	Collin Morris | WI	Lauren Lambert
Other/Unknown Affiliation (3)	Emily Sparer-Fine | MA	Lesliann Helmus
Amanda Deering	Fran Lendacki |Chicago	Luis Hernandez |GDIT
David Kuhar	Greg Fowler | SC	Marie Haring Sweeney
Karla Armenti |UNH	Hanna Chakoian | MN	Matt Stuckey
	Jacki Griffin | NYS	Myrna del Mar Gonzalez | GDIT
	Jim Sainsbury | IN	Nicole Lindsey
	Jordan M | NH	Peter Boersma
	Karene Thomas-Watts | CO	Robert Pratt
	Kayleigh Nerhood | GA	Sandy Roush
	Kelsie Ostergaard | DC	Stacey Marovich
	Ketki Patel | TX	Tashina Robinson
	Kyra Kofodimos | IN	Yan Yuan
	L Espinosa | IN	
	Margret Watkins | WV	
	Marija Borjan | NJ	
	Nelson Lu | Chicago	
	Sequoia | DE	

	Vajeera Dorabawila | NY	

**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*):

Meeting Ground Rules and Updates

- 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.

Data Element-Specific Teams (DEST)

- DEST member expectations – members are not expected to write documents, lead meetings, or prepare slides. The goal is to keep members engaged in conversations around content during and between meetings.
- There is an expectation of active participation in polling, document review, discussion, and other input during and between meetings.
- **[Action Requested]** Please sign up for upcoming data elements on the DEST Sign-Up here.
- After SOGI or Industry and Occupation wrap up, the next data elements will be
    - Exposure/Risk/Underlying Conditions
    - Immunization History

Core Data: In Scope for DSWG

- Focus on core data for surveillance for transmission to CDC within selected data classes.
- We will work to identify standard definitions and value sets within that definition of core data.

Core Data FAQ

- Identification as core does not mean it is required to fill in or send an answer for every core data for every case; it is okay to leave it blank or unknown.
- Data, elements, and data classes should be considered core if they are relevant for many conditions.
- A good way of thinking of this is if we were to create a GenV3 MMG, would it make sense to include that data element here as core?

Industry and Occupation Discussion

Potential Core Data Elements

- This is the list of data elements the group began discussing in last month's meeting.
- The first two are the current industry and occupation data elements in the standard template across NDSS message mapping guides.
- In October, the group voted with more than 80% consensus that these 2 data elements should be considered core data. The group also agreed that the value sets should be switched from the census code system to the NAICS and SOCS code systems.

- Currently, these elements are defined just as cases, current industry or cases, current occupation. The main discussion of these two elements last month centered on this definition and whether the current timeframe should be tied to a description like the time of the event or a date value like the calculated case counting date (CCCD). In previous work, this group defined the CCCD variable for case surveillance.
- The third row in this table includes potential wording for a template for collecting condition-specific information on a particular occupation, such as food handler or healthcare worker. As the blank line indicates, this data element would not exist the same across cases. Still, the group could recommend a general format for the data element and value set when asking about a specific occupation for a condition. Valuable information was obtained on this template idea in the between-meeting poll. However, a decision has been made to pause discussion on the template as it would ultimately not meet the scope of the current work with defining core data elements. The information gathered will be considered in the future data class work surrounding exposure and risk factor information, which will be discussed after the SOGI data elements DEST wrap-up.

CCCD

- CCCD is the earliest available 6 component dates associated with a case. These dates are listed on the right side of the slide and encompass a range of dates between symptom onset and the date the case is entered or recorded.
- The important feature of the CCCD is that it's possible to calculate a value for all cases, even if all components are unavailable.

Current I/O

- The drawbacks discussed in the last meeting to using CCCD to define the time of the event for current industry and occupation are that it might change as more dates are documented or updated during a case investigation. It might not allow capturing all work-related information needed for conditions with a longer exposure period. Using a phrase that only specifies the time of the event and the definition has less specificity may reduce clarity during the collection or interpretation of data. Still, collecting all information relevant to a particular condition could provide flexibility.

Between Meeting Poll

- 19 responses were received. Thank you to all who participated. It is essential to understand why this information is being collected.
- Today, we'll discuss the top three themes of the responses, which included
    - To capture risk for exposure
    - Identify relationships between work/disease
    - Create occupation-specific messaging to prevent exposures
- Most respondents agreed that if using "time of event," we need to define "time of event" better within the interpretation notes.
- How respondents describe the time of the event is split between a more general time of exposure to the causative agent versus aligning this definition with the CCCD.
- Respondents were then asked to choose their one preferred definition for current industry and occupation, and the highest vote was to define it as at the time of the event and then include better interpretation notes with more explanation about what the time of the event is without associating it with a specific case date.

- When asked the same question, but for respondents to select all the definitions they can live with for these core data elements, most respondents selected the same response. It was encouraging to see that some respondents could live with all definitions. So, while this conversation surrounding the definition is still meaningful, we're not as far from a consensus decision as it might seem.

Discussion

- Today, we will continue this definition discussion and use input from the poll. The focus will be on a potential definition of industry or occupation at the time of the event, with some interpretation notes. There is a suggested language for those interpretation notes on this slide, which is the time of the event for industry or occupation. It should capture all industries or occupations where the case worked during their likely exposure period through illness, onset, or diagnosis.
- Are these notes appropriate or sufficient, or does more need to be added?
- This definition also implies that these core data elements must be able to repeat for a case to capture multiple instances, which we'll also vote on after the discussion. The group was invited to discuss if anyone had any thoughts on these types of interpretation notes in this definition.
- Discussion:
    - Marie Haring Sweeny (CDC) commented that they are not opposed to this but are unsure when talking about current industry and occupation that people might think this is what you are doing today but in terms of what is happening for an infectious disease. The industry for which a person might have had the exposure or date of diagnosis makes sense. Further, those things are already there when we have created a repeating element or template for the current industry and occupation. “Current” is okay because it opposes the usual, which generally means the longest-held job. She considered “current” as occurring when the person was exposed, or the event occurred if they had a short latency.
    - Sam (JMC) noted that this and the standard template are essential parts of the discussion for today. Those should be called “current” industry or occupation, but if the group determines to define it should be “industry at the time of event,” then does it need to be called “current industry at the time of event, or would it be better captured just as industry or occupation for these data element names for case surveillance.
    - Katie Dodd (CDC) wanted to add an example of silicosis, typically a disease of longer latency, as an example for this discussion. So, if we're capturing silica dust exposure at the time of the event, that would potentially be an industry or occupation that occurred decades ago, so it wouldn't necessarily be their current industry or occupation. That is an instance where some confusion could arise.
    - Lesliann Helmus (CDC) noted that what we're trying to capture here are data elements we would capture for every event or that we're encouraging people to submit for every event. A disease could also ask additional questions if they were relevant. So, for silicosis, a person would want to ask about the occupation at the time of the event that would be relevant to that kind of exposure. We're trying to capture the “current” as the common element across diseases and situations. It would be a generic question, but it doesn't preclude a more detailed question or an additional question being asked for a disease-specific situation.
    - Cher Levenson (WA) commented that putting the time of the event into the data element description may confuse people and only focus on whether an individual knows precisely when that exposure was. That might reduce your data collection, which would be a concern because you're tying it to some knowledge about when the exposure occurred.

- Keteki Patel (TX) agreed that if we are looking at any chronic disease conditions, such as silicosis, asbestosis, or many others that we may want to keep track of from our health surveillance standpoint, we need more room to ask for multiple jobs or that clarification of current versus the time of job at the time of initial exposure. One other aspect I am unclear about, even when taking the poll, is whether CSTE or CDC is recommending these data fields into the main reporting system after emergency events only or just in general. From our agency standpoint, they are in discussions with ID folks and others, and there was a sense of hesitation because they had to incorporate it into NEDSS or whatever system that is being used for pushing the data for all of these reportable conditions to CDC using that system. The issue was how many data fields we would have to create to capture multiple occupations and industries for an individual. That piece needs to be clarified for a condition and how that will work, not just for one State but others. Also, a script or a queue for how to ask or pursue this question depends on whether this is for an acute event versus a chronic exposure event.
- Sam (JMC) reminded the group that the goal is to define these potential core data elements for surveillance across all the cases and identify the information that should be collected. If we're trying to collect industry and occupation to identify potential sources of exposure, how should we define what we're collecting in these industry and occupation fields? What if we're tying them to a time of event, as in the case we're reporting? How are we defining the time of the event? What is the time of the event? Is it a date? Is it an exposure period? How can we define this?
- Lesliann Helmus (CDC) wants to bring this back to the concept of looking at this almost as a demographic element. We want the current industry or occupation because we're trying to understand what was happening during exposure. We want the current industry or occupation to use that data element that is characteristically asked. It will likely be in a medical record and will be relevant where we can determine what they were doing at the time of exposure for acute exposure. We want to orient people to that. But still, the current industry is the current industry, and we could have a comment that says please indicate the occupation industry at the time of exposure or occupations and industries at the time of exposure, or likely exposure, etc. A recommendation was made to not complicate the name by adding some of those details. There is a concern about tying this to the time of exposure. We are looking for relevant information about exposure, and if somebody stops working because of the exposure, we don't want to say they are unemployed now. We want to tie into what they were doing at the time. It was felt this was becoming too complicated with a need to focus on the current industry and occupation.
- Sam (JMC) noted that the issue is that conditions being reported in NDSS exist conditions with chronic long-term exposure periods and short exposure periods. If this works in both cases, how can we define the current industry at the time of the event for those long- and short-term exposure periods?
- Lesliann Helmus (CDC) commented that I don't think we can mix those concepts. The current industry is not the industry at the time of exposure whenever it occurred. The current industry is current, and industry and occupation are mutual or historical, but if you got it 20 years ago, that's not current, and just because you were exposed doesn't mean it's current. We need to keep these concepts clean and stick with the standard definitions. So I think what we're trying to tell somebody with this caveat, or the explanation, is, you know, if they got exposed two weeks ago, or if they got exposed a month ago, we're looking for that occupation or those occupations. I would prefer we not try to tie it to the time of the event if it's going to get this complicated because we should not be using the current industry to get exposures that are not from their current job.

- Roberth Pratt noted that they have struggled in TB over the last 30 years to try to identify something in this category. And every time we've come out with a new data collection form, we have shifted what occupation means. At one time, I thought that occupation or industry with some caveat on length of time made sense, but I am coming around to current being current. We have separated it into questions: What is your current industry or occupation? Another one specific to TB, which could have a long incubation period, includes questions like, have you ever worked as a healthcare worker or in a particular industry? There are limitations on the particulars there, like basically health care worker or correctional worker, or somewhere where you would be either exposed to individuals, or you would be in a congregate setting where you are most likely to have an opportunity to be exposed.
- Marie Haring Sweeney (CDC) agreed with Leslie Ann that this is getting too complicated. She liked Robert's comment about putting in disease-specific questions relative to the job. We have been trying to get work into being a core demographic for an NNDSS for a long time. Agrees that we need to have definitions and thinks that current is today. Then, if we must think about a disease like TB with a long incubation period and a long latency, we talk about the disease-specific risks and exposures to make it very clear when we ask about exposure. If you wanted to know about a person's longest-held job, you would ask them about their usual occupation. That is something we could add. I'm not sure that that is the core data element. I think the core is current. When we think about infectious diseases that are included in NNDSS. They have relatively short incubation periods. People do not change jobs often; some do, but many do not. I think the current job is appropriate.
- Chris Dumas noted agreement with what everyone's saying. They added that when you're saying current if, it had in parentheses within the last three months or the last year. This gives a very clear delimiter in the verbiage because if you say current just as the word current without any definition, you will not get a clear standard answer. But if you say current with a clear timeframe, it will help people respond correctly and consistently.
- Sam (JMC) noted that for the core definition, the group would like to see the current industry and occupation not at the time of the event or at the CCCD but just the current industry and occupation. This should likely be used in a repeating block.

Polling

Poll 1: I/O Core Definition

- Question: ***This question was updated in the meeting based on discussion as the following***
  Cases' "current industry" and cases' "current occupation"
- Results
  - n = 39
  - 22 (Yes - support this definition)
  - 12 (Yes - Can live with the definition)
  - 5 (No – Needs different definition)
  - As this is over the 80% threshold, the group will move forward with these as the definitions.

Poll 2: Repeatable I/O

- Question: Should the core Industry and Occupation data elements repeat to allow for capture of multiple values for industry and occupation?

- Results
    - n = 39
    - 35 (Yes)
    - 4 (No)
    - As this is over the 90% threshold, the group will proceed.

Discussion

- Sam (JMC), in reviewing information for our future exposures data class, JMC came across this data element from the carbon monoxide MMG that asks if the patient's CO exposure is related to their current occupation. This felt like an essential idea in this group, as multiple industry and occupation instances can be present for a case. It might also be helpful to have an indicator variable that can signify if a case investigator has noted that a particular job is potentially associated with exposure to the condition. A discussion was opened on whether this should be considered part of core surveillance or what wording could be used.
- Discussion
    - Marie Haring Sweeney (CDC) would love to see a variable like this across all conditions and was curious to see thoughts on this. One question: who would make this determination as related to occupation?
    - Sam (JMC) noted that this would not exist in the EHR but would need to be filled in by the case investigator.
    - Cher Levenson (WA) disagreed with Marie only because of a concern that one of the issues we have with collecting I/O is getting people to understand the usefulness of collecting the information and the existing assumptions about whether something is work-related. This question would be asked after you've gotten all your data, except for in cases where it's an obvious link. This is something that must be thought about more.
    - Marie Haring Sweeney (CDC) yes, when I described it, I was concerned about who would make this determination. You're right; it would have to be after the industry and occupation were collected and somebody decided. On death certificates, we have a check box that says work-related. However, I'm not sure who is making that determination. But it's beneficial to have that information. It might cause some confusion, but having it as a core variable with the I/O and employment status would be helpful.
    - Gillian Haney (CSTE) noted we are talking about the utility of this question versus who would be responsible for making that determination, and those are two very different issues. It was felt the utility of this question is important, and epidemiologists make these types of decisions all the time. It seems the issues are being conflated here.
    - Sam (JMC) agreed the focus of the discussion should be if this should be added as a core data element.
    - Lesliann Helmus (CDC) noted that it will be related to a specific occupation if we ask this. So, it would need to be asked for every job, becoming part of the occupation and industry block. There was concern about adding it in there as part of the block. We must keep in mind that it will be a question asked of every job that's entered, and maybe for most people, it's only one job, but it will become part of a repeating block.

- Sam (JMC) noted this is a good point. This question may have a slightly different utility. If we were going to define industry and occupation as everything throughout an exposure versus current industry and occupation, then is the importance of this question different based on those kinds of potential ways to collect industry and occupation?
- Lesliann Helmus (CDC) commented that even for current occupations, somebody may have multiple occupations. We must understand that we're making it part of that repeating block, which must be handled that way. There was a concern about adding too many questions, but I can see where it would have utility, particularly if somebody does have multiple jobs. It would be helpful to know.
- Marie Haring Sweeney (CDC) noted that less than 10% of the workforce has multiple jobs, so it is not expected to be a big lift. If they have a current job, it tells investigators and future analysts about a particular disease we may need. It may want to look a little harder at whether the exposure was related to one or more jobs a person was holding. I don't know if it's going to have to be explicit. Do you want to have it in the template, or would you have it as a separate question about the condition? Not with every job, but with that with the condition itself. So, age, race, gender, and is the carbon monoxide exposure related to their current occupation when you have three current occupations? So, it gives a little more granularity to the ability to look at the cause of this fatality or nonfatal poisoning.
- Lesliann Helmus (CDC) clarified that the wording would say something like is the patient's exposure related to their occupation?
- Marie Haring Sweeney (CDC) confirmed Lesliann's assumption.
- Sam (JMC) confirmed that this would no longer be part of the repeating block to combine with each industry and occupation question but a separate question.
- Marie Haring Sweeney (CDC) commented that it would likely be a separate question.
- Gillian Haney (CSTE) wondered if this discussion is more appropriate for future exposure and risk discussions.
- Lesliann Helmus (CDC) agreed that it may be something to table for those talks.
- Sam (JMC) agreed this would be better for those exposure/risk discussions.
- **ACTION: JMC noted that a data element is to be considered for the future exposure/risk/underlying conditions; DEST will be the example of the carbon monoxide MMG to input whether the patient's exposure was related to their current occupation.**

Poll 4: Exposure Related to Occupation

- Note: Sam (JMC) moved forward with voting on this data element and stated that if individuals prefer this to go into risk/exposure, repeating block to vote no today. If you like that this question goes within a repeating block for industry and occupation, vote yes.
- Question: As part of the I/O data class, should a data element of "Occupation Related to Exposure" be included as a core data element?
- Results
  - n = 31
  - 15 (Yes – Indicator variable that the patient's exposure for this condition is potentially related to this occupation (Y/N/U)
  - 2 (Yes – other definition)
  - 14 (No – Not Core)
  - The votes are split.
    - *CDC and CSTE noted that given the split decision, this should go with the Exposure/Risk/Underlying Conditions conversations.*

Potential Data Elements from eCR/MMGs

- For the last part of the call, it was necessary to point out other variables in the background research. These include items from electronic case reporting "usual industry" and "usual occupation," defined as where the case has worked for the longest time, as well as Employment status, which captures if the person is employed, unemployed, or not in the labor force. An additional element in multiple condition-specific MMGs ascertains whether or not the person was a healthcare worker.

Poll 5: Additional I/O Elements

- The poll will work to see if any other data elements on this slide need to be discussed as potential core data elements from an industry and occupation standpoint in case-based surveillance data transmission to the CDC.
- Question: Are there any other industry or occupation-related variables that need to be added as recommendations for core data for surveillance
- Results
    - n = 26
    - 3 (Yes)
    - 23 (No)
    - No, this DEST does not need to discuss additional data elements for Industry and Occupation.

Next Steps

- This is the last Industry and Occupation meeting. JMC will work with CSTE and Co-Chairs to develop recommendations and share them with DEST when ready for input.
- The next full DSWG meeting is scheduled for December 8, 2023.
- The next DEST meetings include two calls for December 15
    - Sexual Orientation and Gender Identity (Call 3) from 12-1 Et
    - Immunization History (Call 1) from 1-2 ET
- Upcoming data elements include Exposures/Risks/Underlying Conditions
- Sign up for upcoming data element-specific teams in the interest form here
- Additional input is always welcomed between meetings from DSWG members on the form located here.