Meeting: CSTE Data Standardization –Laboratory -Data Element-Specific Team (DEST) Meeting
Date: August 9, 2024
Attendees: 66

CSTE (1)	PHAs (29)	CDC (31)
Meredith Eddy	DSWG Co-Chairs	Angie Agunloye
JMC (3)	Laura Erhart | AZ	Anne Peruski
Erin Kough	DEST Members	Ayana Perkins
Laura Lane	Andria Apostolou | IHS	Carmen Pugh
Natalie Raketich	Anna Kocharian | WI	Danielle Tack
	Beth Isaac | NYC	Dave Jones
Unknown (2)	Brianna Rossi | VA	Ellen Martinson
Luther Topico	Charlotte Picard | IL	Fatima S.
Yasemin Kaynas	Dave DiCesare | NY	Itir Cole
	Debra Rivera | NM	Jennifer Huang
	EJ Klingler | NYC	John Gerstle
	Elizabeth Lando-King | MN	Lauren Korhonen
	Erica Ruud | NV	Lesliann Helmus
	Erin Doyle | NJ	Lynda Rowe
	Gabriela Escutia | OR	Manjula Gama Ralalage
	Hanna Chakoian | MN	Marion Rice
	Heather Elder | MA	Meeyoung Park
	Jacob Chavez | LA	Melissa Pagaoa
	Karen Arrowood | TX	Myrna del Mar González | GDIT
	Kayleigh Nerhood | PA	Nicole Gallaway
	Lunden | IN	Nicolette Bestul
	Mario Beltran | MD	Noreen Kloc
	Matt Kaneshiro | WI	Peter Daly
	Melissa Kealey | CA	Rachelle Oseni
	Misty Johnson | WI	Renee Boehmert
	Morrow Toomey | AK	Sandra Roush
	Nancy Barrett | CT	Sara Johnston
	Rob Laing | OR	Shawn Arshad | GDIT
	Shawna Stuck | GA	Shelby Lyons
	Sophia Cantor | WA	Sophia Kazakova
	Stella Tsai | NJ	Tricia Aden

**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*):

Laboratory

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
    - Reporting Source – first call 8/16
    - Social Determinants of Health – first call 9/20
    - Race/Ethnicity/Tribal Affiliation
- 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

- Out for DSWG-level review:
    - Exposures, Risk Factors, and Underlying Conditions
    - Signs, Symptoms, and Clinical Manifestations
- Coming Soon for DSWG-level review
    - Disability
- Coming soon for DEST-level review
    - Identifiers
    - Medications
- *Note: All drafts released for DEST/DSWG review are drafts; the resulting recommendations passed to CSTE for final steps have not been made available. Please do not pass on draft documents as final products.*

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.
- This DSWG Post-Meeting Input Form is available at any time to provide additional input

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.

Laboratory Data Elements (Part 1)

Laboratory in two parts…

- The overall goal for this data class is to discuss a Standardized Laboratory Reporting Group of data elements. This discussion will be covered across at least two calls.
- Today, the following data elements will be discussed: Test Performed and Test Results, including Units of Measure and Reference Range
- If time allows, there will be a discussion of Specimen Type and Specimen Source Site.

Laboratory: Current State (Lyme/TBRD v1 MMG)

- Currently, laboratory information collected in NNDSS case notifications through message mapping guides (or MMGs) is sent in some variation of what's called an "Epidemiology Laboratory Repeating Group," such as the Lyme/Tick-borne rickettsial disease one shown on the slide.
- These repeating groups ask for a test type with an associated result. The set of allowable test type values varies by MMG and can include LOINC codes, CDC codes, or both.
- These Test Type Value sets can range from as few as eight to hundreds of codes.
- Test results are separated into three categories:
    - interpretive (such as positive, negative, not done),
    - coded organism (such as Borrelia burgdorferi or Borrelia mayonii), and
    - quantitative
- Each of the three result categories is captured as a separate data element, and each MMG includes 2 or 3.
- Each repeating group captures a variety of additional data elements.
    - Most capture specimen type, specimen source, and specimen collection date.
    - Additional data elements captured include performing laboratory type, test method, serotype, genotype, specimen ID, and whether the specimen was sent to CDC for testing.
- Currently, each Epidemiology Laboratory Repeating Group is condition-specific. In this example, the test type value set consists of 24 LOINC codes with an option to specify "other" test types. There are separate data elements for interpretive, coded, and quantitative results. In addition to test type and result, specimen type and specimen collection date/time are included.

Laboratory: Current State (COVID-19 V1.2 MMG)

- This slide shows the Epidemiology Laboratory Repeating Group from the Covid-19 MMG.
- The test type value set for COVID includes 93 LOINC and CDC values.
- There are two test results data elements, one for interpretive results and the other for quantitative test results.
- In addition to these data elements, specimen source, specimen collection date, performing lab specimen ID, performing lab type, and WGS ID Number are all captured.

Laboratory: Current State (CDC Minimal Data Necessary for Caste Notifications and other lab-related data elements)

- Anyone who attends other CSTE workgroup calls may be hearing about the *Minimal Data Necessary (MDN)* for Public Emergency Response.

    - In November 2022, the CDC Advisory Committee to the Director Data and Surveillance Workgroup recommended defining the Minimal Data Necessary (MDN).
    - While lab data are handled as their domain within the MDN, lab-related data elements are also included as MDN for case notification recommendations, which are shown here.
- These lab-related fields would be included in case notifications at the beginning of a public health response. These data elements would be included in case notifications for all cases reported to the CDC and other epidemiological information. These data elements transmit information on tests associated with the case, including information on the test name, test result, and specimen type.
- In addition to information about individual test results, the MDN for case notification includes the Earliest Specimen Collection Date Associated with a Positive Lab Result and the Earliest Result Date of a Positive Lab Result.

Proposed Data Elements: Part 1

- Here are the proposed data elements to discuss first today.
    - Test Performed – a code for the lab test or analysis performed on the specimen in question.
    - Test Result – a universal test result data element allowing interpretive, coded organism, or quantitative results to be sent in the same field. In cases of HL7, the test type would be specified in the OBX-2 segment.
- In addition to these foundational data elements, there are also the two sub-elements of the Test Result, which are
    - Units of Measure
    - Reference Range
    - these data elements establish those variables when applicable to the test result.

Proposed Data Elements: Future Call

- A note of what will be covered in the future.
    - After we discuss the four elements outlined on the previous slide, we will talk about
        - Specimen Type and
        - Specimen Source Site
    - Also included in future discussions are
        - Test Method
        - Specimen Collection Date
        - Date/Time of Lab Result, and
        - Whole Genome Sequencing ID Number

Proposed Data Element: Test Performed

- Knowing what will be discussed in the future, the group will use the rest of today's call to consider the first four elements introduced, starting with Test Performed.
- This data element underpins the Standardized Laboratory group being proposed here. The definition is hopefully straightforward: "Code for the lab test or analysis that was performed on the specimen."

- The primary change to this data element from what currently exists in the case notification Epidemiology Laboratory Repeating groups is that the value set would not be pre-determined by individual CDC program areas, and instead, any valid LOINC would be accepted. There would also be no LOCAL codes accepted, although there is an Other, specify option.

Proposed Data Element: Test Performed (continued)

Discussion

- Erin requested comments or concerns about the definition proposed or comments on the changes to the value set.
    - Nancy Barrett requested clarification for the use of "other." She believed it's typically used when there's a new test and that there is usually a collaboration between APHL and terminologists, who would either find a link for us or submit a request for a temporary link. This way, 'Other' would not need to be used. Concern was expressed that using others could lead to non-standard values being entered.
        - Erin Kough clarified that what was described is the preferred route. However, the CDC knows the need to leave an option open.
        - Misty Johnson noted that in Wisconsin, they often use 'Other' because when ELR data comes in, it parses and fills in their form nicely. However, it's recognized as text since it goes through the parser. Even if there is a valid LOINC code, they can't map it as a coded element because the system treats it as text. So, they usually set the coded element to 'Other' and include the verbiage with the LOINC code. Misty advocated for keeping it this way, though the concerns were acknowledged.
- Erin guided the discussion to the value set. First, regarding the lack of local codes, are there any advantages or disadvantages to allowing any valid LOINC code for the jurisdiction? Would that be a problem or a benefit?
    - Misty Johson commented that they don't currently have to load the value set or vocabulary for something like this because they always use 'Other' and put it in the 5.9 position. However, there is a situation with result units—that vocabulary is massive. They've had trouble using it and rarely receive units, so they don't usually send them. When the vocabulary gets that large, it tends to be problematic, and a concern was raised regarding all the different LOINC codes.
        - Erin noted that test results and tests performed will be part of the ongoing conversation.
    - Nancy Barrett shared their experience with TB HL7 messaging based on the MMG. Although good LOINC codes were available— and had been requested from APHL—the CDC still created a batch of PHC codes for the molecular tests. As a result, they had to configure our system to recognize these PHC codes as full LOINC codes, which raised some concerns. Nancy contacted an APHL terminologist to obtain a complete and appropriate set of LOINC codes for these values. This experience reinforced her belief that we should avoid using local codes whenever possible, as it complicates things. They had to add extra coding to accommodate local codes when standard codes were available. This situation caused a lot of issues, and it took us a while to find a workaround that allowed us to send a local code as a LOINC code, ensuring the gene name was correctly recorded for those TB molecular susceptibility test results.
    - Jennifer Huang wondered if when we define 'test performed,' does that refer to the brand name level, the method level, or something else? She was not entirely clear on this from the LOINC side, as it seems it could vary depending on the test. We deal with several tests, so is there a specific definition for what would be considered a test performed?

        - Erin Kough noted that 'test performed' refers to the specific test conducted. If you're asking about a panel, the 'test performed' would be the individual component of the panel that was performed, not the panel name itself. So, it's not necessarily the ordered test, which might be the panel name, but rather the specific test that resulted from it.

Proposed Data Element: Test Result

- This field would be able to transmit all types of result data.
    - In the case of HL7 messaging, the data type would be specified in the OBX-2.
    - The values set for coded test results would include SCT codes from organism hierarchy to capture organism, any SCT code found in other hierarchies, and appliable APHL SCT Extension/PLR codes.
    - There would also be an Other, specify option.
    - There would be no lists of results for the CDC program area specified.

Proposed Data Element: Test Result (continued)

Discussion

- Erin Kough opened the discussion by asking about the advantages and disadvantages of sending all types of results in one data element. I would ask if anyone sees this as a significant barrier at the jurisdictional level. Specifically, does anyone see the idea of sending both coded and quantitative results through the same data element as potentially problematic?
    - Misty Johnson noted concerns, though she wouldn't necessarily call it a barrier. Over the last few message mapping guides we've onboarded, starting with COVID-19, we've been able to extract data from our ELR. While there are some quirks, such as everything being in text format, we handle qualitative and quantitative results separately in our lab widget. She would prefer to keep these results distinct. For instance, in preparing for this call, I looked at Lyme, tick-borne diseases, and TB, which all have separate fields for quantitative and qualitative results (e.g., INV 291 for numbers and LAB 628 for qualitative results). She didn't recall seeing an HL7 line where we constantly switch the OBX-2 value between coded and structured or numeric. She was unsure how challenging that would be, but she preferred to keep them separate.
    - Nancy Barret commented that she recently participated in a small group led by Peter Boersma to discuss laboratory reporting and message mapping guides. Along with colleagues like Misty and others, we considered how to handle ELR data. The consensus was that each OBX segment should define a result based on its point code, with separate OBX fields for different results. For example, suppose we receive an antibody result for hepatitis C, followed by a PCR result with a quantitative measure. In that case, these will come in as separate OBX segments, even from the same specimen. We prefer to handle and send data the same way we receive it. So, if we receive results as separate OBX segments, we'd like to send them back to the CDC in the same format. This approach helps with epidemiological interpretation and case classification, allowing for clear reporting of qualitative and quantitative results. Ideally, there would be one line for qualitative results and another for quantitative results if needed, mirroring how we receive and store the data.
        - Erin Kough replied that she understands Peter worked with a subset of states. However, at least one other state does not automatically map lab results to the correct NNDSS message mapping spots. She understands that Wisconsin isn't alone in this issue of not automatically mapping HL7 results from ELR. Hanna Chakoian confirmed that MN is another state like this. She called attention to a chat comment mentioning that a jurisdiction receives both coded and quantitative results in the same

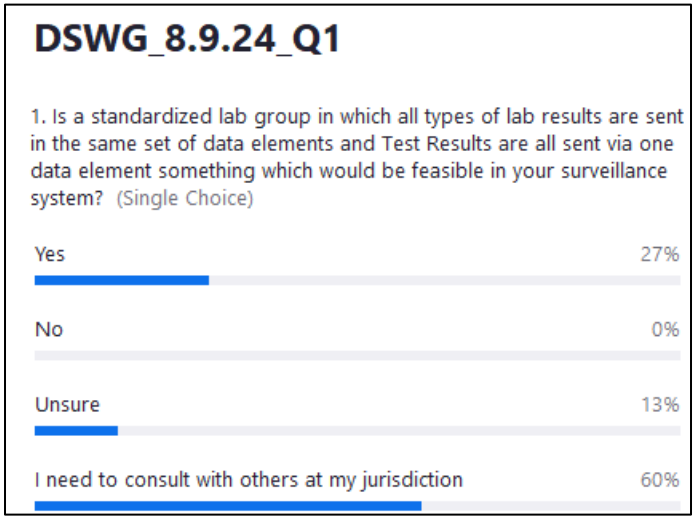
message and would like to reduce the workload. Could you clarify if consolidating these into a single test result would help reduce the work?

- Erica Rudd replied that standardizing results to one type or the other would help resolve their issue. Their system currently struggles with interpreting OBX-2 because sometimes it's a SNOMED code, and other times it's a different format, like 4.2. This inconsistency requires us to create workarounds for coding. Standardizing to a single format would simplify this process. However, this issue might be specific to certain SNOMED codes and may not apply to all.

- Erin asked the group if any major barriers were associated with local codes not being accepted as part of the test results value set.
  - Hanna Chakoian commented that it would be problematic if Minnesota sent their raw data directly because they have some local codes. Some entries might indicate 'no code given,' while others are more specific. In such cases, they must set up a process where someone reviews and translates these into appropriate codes. It's not ideal, but they could configure it to make it work.

Poll 1 (n= 15) – Standardized Lab Group

- The first poll is specifically for jurisdictional partners. We would like to know if a standardized lab group where all lab results are sent in one repeating group, including the test results all being sent as one data element, would be possible in your surveillance system. It was noted that if participants are epidemiologists and are not as confident about mapping data elements from their surveillance system to a back-end transmission format, please do not select the fourth option.



Discussion

- Karen Arrowood raised an important question asking how standard results from point-of-care tests, which are manually entered and might not have LOINC or SNOMED codes, would be transmitted.

        - Angie Agunloye requested clarification: are these results coming in as text, or do they have some local codes for point-of-care tests?
        - Karen Arrowood clarified that these would typically be the text that would be input into the system and would not have any codes attached.
        - Angie Agunloye replied that they're still allowing flexibility by using 'Other' for some point-of-care tests that might not have standardized codes. Given that 'Other' is permitted, which accommodates free text in OBX-5.9, and considering that they are receiving these results as text values, would it be feasible to transmit them within OBX-5.9?
        - Erin Kough noted that this seems important to consider.

Proposed Data Element: Units of Measure/Reference Range

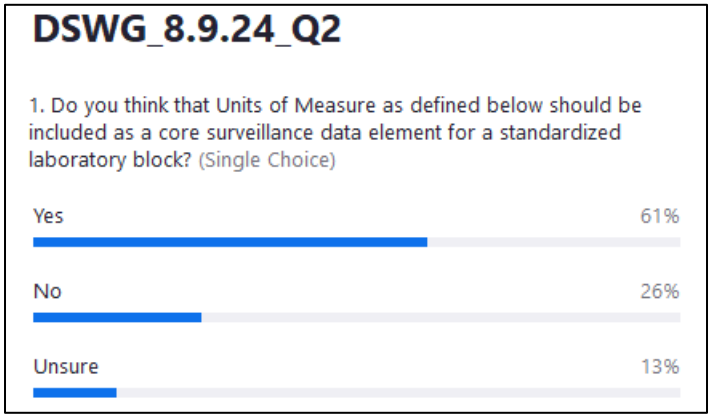
- The discussion flowed into Units of Measure, which would be sent via OBX-6 in an HL7 format, and Reference Range, which would be sent via OBX-7.

Discussion

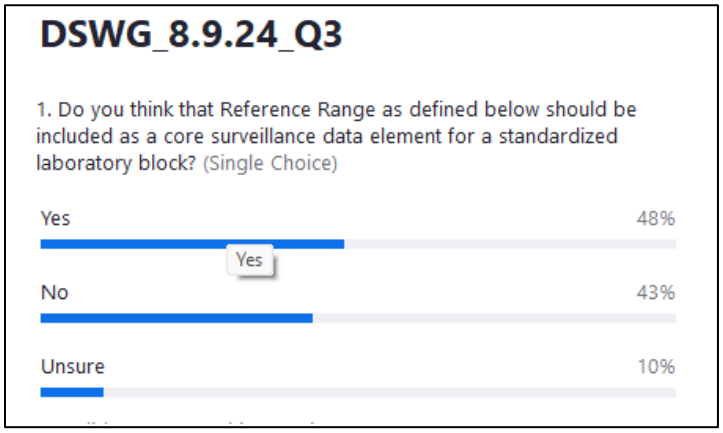
- Erin Kough noted that some concerns were mentioned earlier regarding these data elements. To clarify, units of measure would be sent as OBX-6 in HL7 format, and reference range would be sent as OBX-7. The units of measure would use a unified code value set, while the reference range is what's reported. We haven't yet discussed whether any data elements should be considered for core surveillance, as core surveillance typically requires lab information of test performed and test result. However, we were asked if these sub-data elements should be included in core surveillance as defined previously. Additionally, feedback was requested on the inclusion of abnormal flags. Finally, are there any concerns or questions about the units of measure and reference range data elements, including their definitions and value sets?
    - Misty Johnson mentioned that everything they receive, except for dates or numbers, should ideally be a coded element. However, if the unit is text (like microliters), and the unit of measure vocabulary is extensive and lacks an 'Other' option, they haven't been able to send it. They don't always receive units in the lab reports. While it's theoretically possible to work on a transformation to match text strings like 'microliters' to the expected units for a particular condition, this hasn't been pursued. If we're considering this as a minimum data set or core element, it would involve a complex text-matching exercise, which is not feasible for them. So, in Wisconsin, they won't provide units of measure. If 'Other' in OBX-5.9 is used, they might be able to work with that.
    - Erin Kough noted that within the chat, Oregon and Louisiana are also facing challenges with this issue, particularly with STI lab data. Nancy mentioned that they collect units based on UCUM but accept them as text in a separate data field in the lab reports. It appears this is a tricky issue for many jurisdictions.
    - Nancy Barret followed up on Misty's questions and wondered what the CDC does with quantitative results and reference ranges. For example, does the CDC use this information if we provide quantitative results with units and a reference range? She knows that specific reference ranges may be used for some tickborne tests to interpret titer levels. Is this information utilized by CDC programs that deal with these tests, or is it more like additional context, such as an abnormal flag, rather than being critical to the result? Any insights on how this data is used would be helpful.
        - Angie Agunloye responded that the information you provided, including the quantitative results and reference ranges, is useful. While narrative details and basic positive or negative results are important, programs like HAI would find the quantitative

results and reference ranges particularly valuable if they are included in the OBX segment. This additional detail helps interpret the data more effectively.
    - There was a question about whether an abnormal flag adds anything to the case information since all results sent should be in the "abnormal" range. However, Hanna Chakoian replied that from their experience with automating COVID results, sometimes the only coded result available is the abnormal flag. There might be additional information in the comments, or there might not be.
- Nicolette Bestul noted that since Rickettsia was mentioned, she wanted to add that for their case definitions, they use a minimum titer of 1:128 for most diseases. They don't typically use the reference range on our end and trust that the states sending us cases have already made their judgments before forwarding the information. Once the data reaches us, we don't rely on the reference range.
- Anne Peruski commented that she works with HIV surveillance, and we use reference ranges primarily for data quality checks, especially for viral load test results. Although we don't use the reference ranges directly in our analysis, they help us identify data entry or processing errors. For example, if the lower limit of quantification is 20 but we receive results of 5, it likely indicates a data error. So, we rely on reference ranges to ensure the accuracy of our data.
- Sara Johnston noted that while this perspective isn't specific to any CDC program, it's important to consider the general utility of data elements like reference ranges. If the reference range is necessary for interpreting results in your system, then the CDC would also benefit from that information. Similarly, if abnormal flags or other data elements are used for interpretation, the CDC would need them to interpret the results accurately. Since the requirements can vary greatly by test and lab, if these tools are valuable for your purposes, they could also be helpful for the CDC.
- Carmen Pugh wanted to clarify a point about HIV and reference ranges. For viral loads, there isn't a traditional reference range; instead, what's reported is the detectable range or the level of detection. Since viral loads are not typically within a normal range for conditions like HIV, there is no reference range in the conventional sense. It's important to distinguish between a reference range and a detectable limit. Currently, there isn't a specific field for indicating the limit of detection in these reports

Poll 2 (n=23) – Units of Measure as Core



Poll 3 (n=21) – Reference Range as Core



Proposed Data Elements: Specimen Type/Source Site

- This is a peek into what is proposed for specimen type and source type – these wouldn't be expected to be populated unless they could be pulled automatically from an ELR, and they would not be expected to be derived from a LOINC code, for example.
- The Specimen Type is defined as the type of specimen used in testing the resulting lab test.

- The SNOMED Specimen hierarchy is the value set for this data element. The Specimen Source Site "Indicates the anatomical source of the specimen tested." The value set would be SNOMED Body Site hierarchy.
- It was noted that this would be discussed more in the second call and that this was an introduction to get participants to start thinking about these concepts.

Next Steps

- Next Laboratory Reporting call on 9/13
- Upcoming team meetings
    - 8/16, 1-2 ET, Reporting Source
    - 9/20, 1-2 ET, first SDOH DEST call
    - Link to DSWG: Data-element Specific Team Interest Sign-Up
- The next large DSWG call will be on August 30, 2024.