Meeting: CSTE Data Standardization –Laboratory -Data Element-Specific Team (DEST) Meeting
Date: September 13, 2024
Attendees: 66

CSTE (3)	PHAs (35)	CDC (23)
Becky Lampkins	DSWG Co-Chairs	Alejandro Pérez
Gillian Haney	Judie Hyun | MD	Angie Agunloye
Meredith Eddy	Laura Erhart | AZ	Chirine Chehab
JMC (4)	Molly Crockett | MA	Danielle Tack
Erin Kough		Gonza Namulanda
Laura Lane	DEST Members	Jennifer Huang
Natalie Raketich	Aaron Bieringer | MN	Josh Bagley | GDIT
Sam Spoto	Andria Apostolou | IHS	Kasey Diebold
	Alina Paegle | UT	Luis Hernandez | GDIT
Unknown (1)	Beth Isaac | NYC	Lynda Rowe
Aileen Duldulao	Brianna Rossi | VA	Manjula Gama Ralalage
	Charlotte Picard | IL	Myrna del Mar González | GDIT
	Deb Loniewski | MI	Nicolette Bestul
	Debra Rivera | NM	Peter Boersma
	Devann Kirkpatrick | TN	Rachelle Oseni
	Ed Hartwick | MI	Renee Boehmert
	Erica Ruud | NV	Sandra Roush
	Genie Tang | CA	Sara Johnston
	Hanna Chakoian | MN	Shawn Arshad | GDIT
	John Satre | IA	Shelby Lyons
	Kacee Redden-Benz| SD	Sonia Singh
	Karen Arrowood | TX	Sophia Kazakova
	Kayleigh Nerhood | GA	Tricia Aden
	Kiara Vicente | AK	
	Kyra Veatch| IN	
	Mario Beltran | MI	
	MDH Elizabeth Lando-King | MN	
	Melissa Kealey | CA	
	Misty Johnson | WI	
	Monique Birmiel | IN	
	Nancy Barrett | CT	
	Nora Kelly | IL	
	Rob Laing | OR	

	Shawna Stuck | GA Suzanne Gibbons-Burgener | WI Theron Huntamer | NV Tim Liao | NYC	

**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 Reporting (Call 2)

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
    - Social Determinants of Health – first call 9/20
    - Race/Ethnicity/Tribal Affiliation - 10/25 on the broader DSWG call
- 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:
    - Identifiers
- Out for DEST-level review
    - 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 2)

Laboratory in two parts…

- The goal of this DEST is to put together a standardized laboratory Repeating group for case notifications for all cases requiring lab information.
- In August, we discussed the four data elements listed here.
    - No polling was done related to Test Type and Test result as they are the underpinning of laboratory information, and a repeating group would be useless without them.
    - We completed a poll for Units of Measure, which had weak support for inclusion, and the group was split on whether Reference Range was useful enough to include in a standardized repeating group.
    - There was some discussion about a Test Result type data element, which would establish what kind of result a test was expected to have, for example, a coded result such as positive or negative. Jurisdictional members of the DEST were asked to weigh in on whether this would be feasible in their systems, but it was decided that the question would be more appropriate to pose to jurisdictional staff with a more technical messaging focus. More follow-ups will be done on this data element in the future.

Laboratory in two parts… (continued)

- Today, we will discuss the rest of the data elements proposed for inclusion in a standardized repeating group for laboratory information transmitted to the CDC in case of notifications.
- We will discuss these data elements in sections starting with Specimen Type and Specimen Source Site.
- We will work through the groups systematically, and if needed, we will go to a third call in October.

Proposed Data Elements: Part 2

- This slide provides more details on the proposed data elements and suggested value sets, which we will discuss today.
- The suggested value sets for specimen Type and Specimen Source Site move away from condition-specific or CDC-specified lists and toward SNOMED hierarchies, which would be able to accept values being ported from an ELR into the case notifications.
- The Performing Laboratory Type will be discussed after the specimen detail data elements. This data element maintains a definition like the ones used in the past, but the value set is simplified to meet the CDC's needs.

Proposed Data Elements: Part 2 (continued)

- This slide details the rest of the data elements. We will discuss the two date/time data elements and the two identification numbers together.

Proposed Data Element: Specimen Type

- Specimen Type is defined as the type of specimen used in testing the resulting lab test.
- The proposed value set is the 2024 SNOMED specimen hierarchy, which allows for varying levels of specificity.
    - Jurisdictions able to pass results through from ELR to case investigation or notification could map the SPM-4 segments to this data element.
    - Jurisdictions unable to do this would be able to map their specimen type options to high-level specimen types such as blood and urine. There would be an “other, please specify option” and a place to send text with it.

- The CDC program would not be able to designate any condition-specific lists of specimen types, and all values within the hierarchy could be received at that end.

Proposed Data Element: Specimen Type (continued for discussion)

- With this definition and value set in mind, the following prompts opened the discussion. Does anyone see any advantages or disadvantages to the all-inclusive value set? Are there any barriers that anyone might be concerned about?

  Discussion

  - Alejandro Pérez asked if the value set should be tied to the 2024 SNOMED version or if it would be better to reference SNOMED codes in general without linking to a specific version. What happens if, in the future, a lab uses retired codes? Would referencing a specific version cause issues over time, such as in 5 or 10 years?
    - Erin Kough responded that there have been issues in the past due to not specifying versions clearly, but I'll let Angie or Peter weigh in. Could it be problematic if someone sends data that isn't in the expected lists?
    - Suzanne Gibbons-Burgener noted that it could say 'current SNOMED specimen codes,' but updating the system can take time. Is there a preference for using codes from the past year, or is there a standard process for adopting new versions when updates are released?
    - Peter Boersma replied that we should ensure we capture new codes by regularly updating the value set to avoid missing any in use elsewhere. The concern about deprecated codes is more complex—if they're still being used and transmitted, we might want to allow them, but our priority should be staying updated with the newest codes.
    - Nancy Barrett noted that the reference is to SNOMED CT, specifically the US Browser edition. She has never heard of it called '2024 SNOMED.' If you look it up, SNOMED CT is the name of the data source, and she hasn’t seen any SNOMED codes deprecated—thousands are available. New preliminary codes can be requested, typically for new organisms, though she hasn’t seen one. There's a process for assigning preliminary numbers, like Regenstreif’s process, but it's slower. In many systems, SNOMED is the primary source for coding. HL7 has an approved table for source and source type, though it isn't the ELR standard. Some states may accept local codes, but we generally stick to SNOMED. However, we might need to allow flexibility for other codes if systems can't send the standard ones, as discussed previously with LOINC codes.
    - Ed Hartwick agreed with Nancy on her points. He liked the idea of including HL7, as it's been used for a long time. We don't want to exclude anyone if it's possible to use. We've struggled with getting hospitals and other sites to send SNOMED specimen codes, so encoding to send to the CDC wouldn't be too difficult for us. However, using it as a pass-through could become more challenging since many sites still can't implement it, even after years of trying.
    - Angie Agunloye wanted to clarify that SNOMED CT, like LOINC or any standardized coding system, has versioning. SNOMED updates occur twice a year—in March and September. The '2024' reference just indicates the current version. On the SNOMED browser, individuals can select different versions, such as March or September, which might show deprecated or new concepts. The reference to '2024' reflects what's current now, but during implementation, we’ll use the version available at that time, like the 2025 March or September version.

        - Nancy Barrett asked if we just say, 'the most recent version' or 'current version,' we can avoid year-specific tagging and the need to constantly update it. Angie mentioned that the newest version is typically presented first unless you choose an older one.
        - Angie Agunloye agreed that would be correct.
    - Erin assessed the chat and wanted to address the discussion about epidemiologists having to choose values from a large value set. No one suggests using the entire SNOMED CT hierarchy in your surveillance system—that would be very time-consuming. You'd either map things on the back end or, if manually entering lab data, have a broad selection like 'blood specimen.' It doesn't need all the detailed options, and the CDC system can accept broader categories, such as 'blood specimen,' without requiring specific subtypes like 'venous blood.'

Proposed Data Element: Specimen Source Site

- Specimen Source Site, which indicates the anatomical source of the specimen tested. The value set is the 2024 SNOMED Body Structure hierarchy, which would be able to accept codes sent in an ELR's SPM-8. As with specimen sources, jurisdictions not pulling ELR data directly into their case notifications could map system selections to related SNOMED codes.
- There are no condition-specific value sets for this data element and all SNOMED codes within this hierarchy. There would be an "other, please specify option" and a place to send text with it.

Proposed Data Element: Specimen Source Site (continued for discussion)

- The discussion was opened with the following prompts: Is this information reliable in labs? Would inclusion in a standardized reporting repeating group be the best place to put it? What are the pros and cons of a broad value set?

    Discussion

    - Ed Hartwick commented that these items should be included as core components because even if you can't always send them, it's important to send them when you have the information. He has also noticed that many sites receive 'unknown' entries because they cannot map items from reference labs.
        - Erin Kough noted agreement and that the question is whether we should keep a space open for this kind of data element. She paused to ask the CDC their thoughts on the plan if there is an 'unknown' entry.
        - Peter Boersma responded that he'd need to check the value set to see if an option for an 'unknown' value could be entered. An 'unknown' entry is more informative than leaving it blank. As Ed mentioned, the key point is to send the information if you have it, with 'unknown' indicating that the data is unavailable.
        - Nancy Barrett commented that she wanted to acknowledge a topic in the chat and elaborate that source site and specimen sources were not usually received, not very often from what was seen. However, it was known that TB looked for both the anatomic site of the disease, which is SNOMED-based, and some local codes found in the PHIN VADS value set. They also included specimen sources such as blood, CSF, and other fluids with SNOMED codes and body structure locations, which may have been collected more frequently than for other diseases. Additionally, labs were asked to enter specimen sources into SPM 4 to simplify mapping, as each disease had specific sets of specimen sources, avoiding the need to pick from a huge list. There

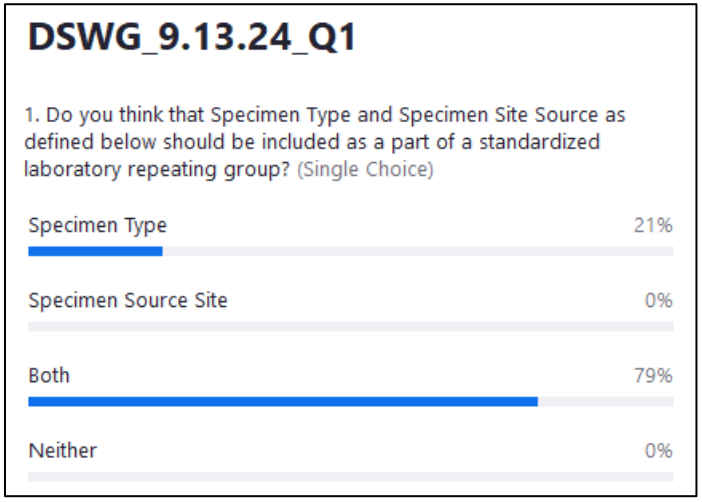
was also a null flavor for 'unknown,' with 'UNK' as the standard null flavor. However, it was usually considered 'unspecified' since labs typically needed some source identified to conduct tests. 'Unspecified' was often used because many labs, like major national commercial labs, did not report specimen sources, which was unusual given that testing requires knowledge of the specimen.

- Erin Kough noted that if a specimen type was sent with a specimen source included, it would also be accepted. For example, a swab of the nasopharynx would be accepted. Regarding labs, it was mentioned that some local codes were found in PHIN VADS, which the SNOMED body structure hierarchy would replace. She also acknowledged a comment in the chat about the CDC prescribing a value set of standard codes. Cross-mapping is required, and she believes using these hierarchies aims to address some of that burden.

- Alejandro Pérez noted in the Division of STD Prevention at the CDC that specimen source was important. If the question was whether to include this as a standard, the preference would have been to vote 'yes' if it was considered core, given that there was no condition-specific version of that value set. However, if it was just a standard to be included in individual condition-specific guides, there would have been the flexibility to add it as needed. He advocated that specimen source is a high priority for them.
    - Sandra Roush agreed with this comment and that the level of detail or specimen source is important for vaccine-preventable diseases.

Polling

Poll 1 (n= 38) – Specimen Type Specimen Site Source



Performing Lab Type: Current State

- Performing Laboratory Type. A quick review of what this looks like in current MMGs and how programs at the CDC use it.
- The table at the top of this slide is an example of Performing Laboratory Type. This data element aims to assign the performing lab to categories of lab types, including CDC, state public health laboratories, commercial labs, and others. Versions of this question exist in 8 MMGs. There are also additional questions related to specific lab types, such as asking if a specimen was sent to CDC when a specimen was sent to CDC and if an isolate was sent to a state public health lab.
- Performing lab type enables the CDC to know what kinds of labs are doing tests for different types of tests, and, in the case of discrepant results, the CDC will generally defer to results from public health laboratories.

Proposed Data Element: Performing Lab Type

- The proposed data element for Performing Lab Type attempts to capture some of the more important nuances of this data element while also simplifying data collection and mapping for jurisdictions.
- The definition is classifying the performing laboratory by its primary function, such as a CDC laboratory, public health laboratory, or other types of laboratories.

Proposed Data Element: Performing Lab Type (continued)

- The value set for this data element has been limited to four options. The current values set options are shown here on the left. Not all options are used for each condition, which uses a version of this data element. On the right is the proposed standardized value set. It is limited to CDC laboratory, state public health laboratory, other, and unknown.

Proposed Data Element: Performing Lab Type (continued for discussion)

- The discussion was opened with the following prompts: Does simplified value set lower barriers to mapping/transmitting the data from jurisdictions to the CDC? Does this data element provide useful information to many CDC program areas?

Discussion

  - Nancy Barret commented that from the jurisdiction side, a simplified list is preferred over trying to categorize complex lab types. Hospital labs are straightforward, but distinguishing other types, like VVD labs, has been challenging due to outdated categorizations. With increasing test result sources beyond hospitals and large commercial labs, managing a CDC laboratory or public health lab list is more feasible. Public health labs are a smaller group than all possible lab types and are easier to handle. A shorter list, if necessary, would be more effective.
  - Erin Kough noted that in the chat, Hannah Chakoian suggested exploring whether non-infectious conditions, such as lead exposure or other environmental factors, might fit into this framework. This could be worth investigating, as it may include aspects not typically considered for infectious diseases.
  - Suzanne Gibbons-Burgener shared that there doesn't seem to be much utility for this at the state or local level. Prioritizing the collection of this information may not be necessary. If all laboratories have an ID tied to more specific information about their lab, that ID could be used to determine the lab type. However, most data come from clinical, commercial, or pathology laboratories, and there are challenges in obtaining results from pathology. It's unclear whether to classify pathology as a separate lab. Currently, this data element isn't used.

        - Ed Hartwick noted agreement with Suzanne. From the state's standpoint, it wasn't something used very often. It was understood why the CDC might need it, but there was concern that it might become too broad to be useful. On the other hand, as Nancy mentioned, mapping is easier with four categories instead of six. Even if it was collected in a more granular form, mapping it up would be relatively straightforward. There was neutrality on the issue due to uncertainty about its use beyond passing along information.
    - Erin Kough noted that in the chat, questions were asked about the purpose of this data element and who values it. She encouraged the CDC program areas that support keeping this element to comment.
        - Daniell Tack responded that her understanding included the original use cases and some sensitivity around passing on information, particularly with testing lab names. For some jurisdictions, there was concern about treating this information similarly to PII to some degree. The idea was to have classifications to address this concern, though she did not see the full context. Understanding the specimen ID coming in was important, especially for CDC's data linkage efforts, particularly when tied to sequencing. Many were familiar with the general challenges—often described as the 'chicken and the egg' problem—related to samples, case patients, and public health labs. As data was consolidated, the complexity sometimes increased exponentially depending on the pathogens. It became more challenging to avoid requesting additional data, as there was often a perception that existing data from testing could be used for linkages. This effort was meant to help interpret the information alongside other lab data.
        - Sandra Roush commented that she was in the National Center for Immunization and Respiratory Diseases. One of the previous slides noted some MMGs (8 or 10) that included the performing lab type with more granular valid values. These values were largely harmonized and identical across each MMG in her center. The team had been asked whether they could manage without that granularity, and it was understood that while it would be easier for jurisdictions to map and understand, some more granular options were challenging and not mutually exclusive. The data element itself was found to be very important because not all labs are the same, and discrepant testing was a frequent issue. It was assumed that in the proposal, non-CDC and non-jurisdictional labs would fall under an 'other' category, which was considered helpful. They would always prefer a public health lab or a CDC lab, acknowledging that the CDC is also public health and a state or local entity. In summary, the data element was useful for deciding the most appropriate path for case classifications and investigatory efforts. However, the less granular valid values were also recognized as being useful.
        - Tricia Aden noted that looking at this value set, it seemed that if there was a decrease in categories, including clinical labs, which should be considered. Some categories seemed too broad to be practical. It was suggested that if the element was retained, there should be a way to include clinical sectors in some capacity.
            - Erin Kough commented that it is challenging because not everyone uses “clinical” similarly. She also noted that conversations in the chat were discussed using CLIA numbers.
            - Danielle Tack replied that while she agreed with comments about using the CLIA number, it was mentioned that not all tests are CLIA-approved. The question was whether that number still applied to the lab, even if a public health laboratory test was used. It was noted that such tests might not directly contribute to clinical care but could aid in

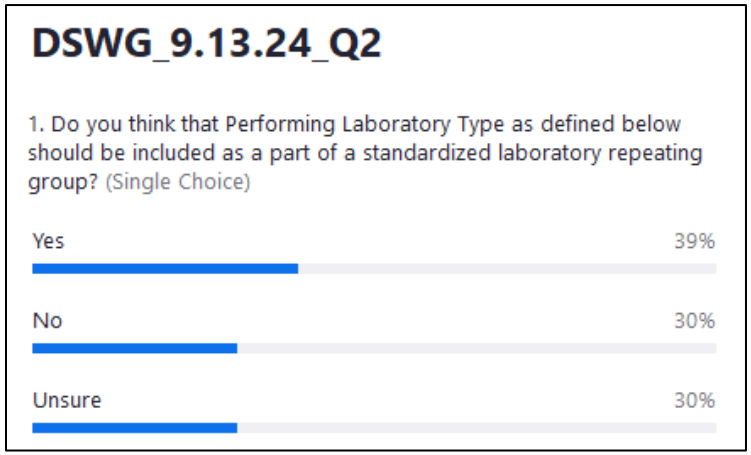
public health actions or responses. Some activities were not intended for clinical treatment but were conducted to help link cases, determine if they were part of an outbreak, or identify information related to transmission patterns. These efforts aim to provide public health support rather than direct clinical care to the patient.

- Nancy Barrett noted that most labs performing emergency use authorizations (EUAs) are CLIA-certified. However, if the discussion involves academic labs conducting specialized testing, such as whole genome sequencing, they might not necessarily be certified labs.
- Danielle Tack noted that her concern was about categories like CLIA. While some public health labs are CLIA-certified, not every test conducted at the CDC or a public health lab is necessarily CLIA-approved. For case surveillance rather than patient reporting, there was curiosity about whether the CLIA number still applied to the lab or if it was inappropriate to use that number when the test was not CLIA-approved. The question concerned the use of the CLIA number. Specifically, for information or reports from public health labs going back to the surveillance system, can the CLIA number still be used, or is it only applicable from a commercial perspective? That was the nuance being considered.
- Sara Johnston commented that CLIA numbers are used to exchange ELR data. Even if a public health lab performs non-CLIA testing for surveillance purposes only, their sender is still associated with their CLIA number when sending ELR records. For CLIA-certified facilities sending surveillance results to their epi group or others, the results are still linked to their CLIA number. The bigger issue might be with sites, such as some COVID testing sites, that do not have CLIA numbers and send data without them.
- Nancy Barret noted that having listened to the earlier discussions, she wondered whether it's valuable to distinguish between CDC labs, public health labs, hospitals, or clinical labs in the messaging. If there's a testing discrepancy, would Sandra's team typically follow up with the location to determine who conducted the test? Is it necessary to include this level of detail in the messaging, or can we proceed with a simplified set for now?
- Jennifer Huang commented that from an antimicrobial susceptibility standpoint, it is helpful to know if the test came from a clinical lab. To her, a clinical lab is any lab that is neither public health nor CDC. What is the value of 'other' and 'unknown' in this context? Is there a space for a 'clinical lab' in the categorization? I apologize if this was covered earlier.
- Sara Johnston noted that there are many ways to classify different types of facilities, and it can be challenging to distinguish between them, especially when their activities overlap. For instance, a hospital lab may also conduct commercial testing, blending the lines between hospital and commercial labs. Additionally, facilities like blood or tissue donation centers may report to public health but are not considered clinical labs. Generally, distinguishing between CDC labs, public health labs, and other labs is useful. In most cases, 'other' would imply a clinical setting rather than a public health facility, allowing for the assumption that the testing followed a more clinical pathway.
- Jennifer Huang responded that there could also be other types of labs in this category. Commercial and clinical labs might be considered similar, but unknowns indicate that the lab type is not identified. Understanding how other

programs use this information is important. Knowing that a clinical lab is involved is valuable because different levels of labs can produce varying AST results. This information helps determine which results should be trusted. Hence, defining and identifying the lab type is beneficial. If the differences were known, it might be possible to infer that a certain method is aligned with public health labs, while another method might be less trusted. However, distinguishing between hospital and commercial labs might not be critical. For categorization, 'other' could include veterinarian labs or various random labs. Combining these groups could lead to merging labs that require action with those that might not need it.

- Erin Kough commented that it is important to note that commercial and clinical categories were likely removed from the value set to simplify mapping for jurisdictions. The goal was to reduce the complexity by allowing CDC and public health labs to define a few specific labs and categorize everything into a separate bucket. Next, a poll will be launched on the proposed value set. The results will guide the next steps, and the discussion on standardized labs will continue in the next meeting.

Poll 2 (n= 33) – Performing Lab Type



Next Steps

- The next Laboratory Reporting (Call #3) will be held on October 11.
- The first Social Determinants of Health call will be held on August 20.
- Upcoming team meetings and additional information can be found here
    - Link to DSWG: Data-element Specific Team Interest Sign-Up
- The next large DSWG call will be on October 25, 2024.