Meeting: CSTE Data Standardization –Laboratory -Data Element-Specific Team (DEST) Meeting
Date: October 11, 2024
Attendees: 57

CSTE (2)	PHAs (29)	CDC (22)
Becky Lampkins	DSWG Co-Chairs	Angie Agunloye
Meredith Eddy	Judie Hyun | MD	Carmen Pugh
JMC (3)	Molly Crockett | MA	Chirine Chehab
Erin Kough		Josh Bagley | GDIT
Laura Lane	DEST Members	Kasey Diebold | CDC
Natalie Raketich	Aaron Bieringer | MN	Kayode Olupinyo
	Alina Paegle | UT	Lauren Korhonen
Other/Unknown (1)	Andria Apostolou | IHS	Luis Hernandez | GDIT
Corrisa Miliander (Mayo Clinic)	Beth Isaac | NYC	Manjula Dharmawardhana
	Brianna Rossi | VA	Marion Rice
	Charlotte Picard | IL	Miriam Van Dyke
	ej klingler | NYC	Myrna del Mar González | GDIT
	Elizabeth Lando-King | MN	Nicolette Bestul
	Erin Doyle | NJ	Noreen Kloc
	Erica Ruud | NM	Alejandro Pérez
	Heather Crawford | OR	Peter Boersma
	Jesse Ellis | MS	Rachelle Oseni
	John Satre | IA	Renee Boehmert
	Kacee Redden-Benz | SD	Sandra Roush
	Karen Arrowood | CDC and TX	Sara Johnston
	Kayleigh Nerhood | GA	Shawn Arshad | GDIT
	Kiara Vicente | AK	Tricia Aden
	Kristen Merrell | MT	
	Kristin Goldesberry | IL	
	Kyra Veatch | IN	
	Lakin Mauch | ND	
	Mario Beltran |MI	
	Melissa Kealey | CA	
	Misty Johnson | WI	
	Nora Kelly | IL	
	Suzanne Gibbons-Burgener | WI	
	Trish McLendon | CA	

**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 3)

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 – second call is scheduled for 10/18
    - Reporting Source – next steps are to be determined
- 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:
    - Revised Industry and Occupation language
- *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 3)

Laboratory in three parts…

- The goal of this DEST is to put together a standardized laboratory Repeating group to be used for case notifications for all cases requiring lab information.
- In August, we discussed the four data elements listed on the slide.

- 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 did 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 on this data element will be coming soon.
- Specimen Type and Specimen Source Site were both supported for inclusion in a standardized laboratory repeating group.
- Performing Laboratory Type was not supported and will not be included in the recommendations.

Laboratory in three parts…(continued)

- We will discuss the rest of the list of data elements proposed for inclusion in a standardized repeating group for laboratory information transmitted to CDC in case notifications.
- We will discuss these data elements in sections starting with Specimen Collection Date/Time and Date/Time of Lab Result and then moving into the two specimen identification number data elements.

Proposed Data Elements: Part 3

- Here are the details for the rest of the data elements.
- We will discuss the two date/time data elements together and the two identification numbers together.

Proposed Data Element: Specimen Collection Date/Time

- The first for discussion is a well-known one, specimen collection date/time. The definition for this data element has been pulled directly from the Data Standardization: Dates of Importance to Public Health Surveillance, Illness (Symptom) Onset Date, Report Dates, Laboratory-Related Dates, (Clinical) Diagnosis Date Brief, which is important to align our work with for purposes of standardization.
- Definition: Date (and time as available) a clinical specimen was collected from the patient for the laboratory testing relevant to the disease or condition being reported to a public health agency.

Discussion

- The discussion was opened with the following prompt: Any discussion on barriers to this data element being transmitted? Or about the utility of the data element?
- *There was space opened for discussion, but no discussion was needed*.

Proposed Data Element: Date/Time of Lab Result

- The next date in question is the Date/Time of the Lab Result, defined as the date (and time as available) the latest result was released for reporting.

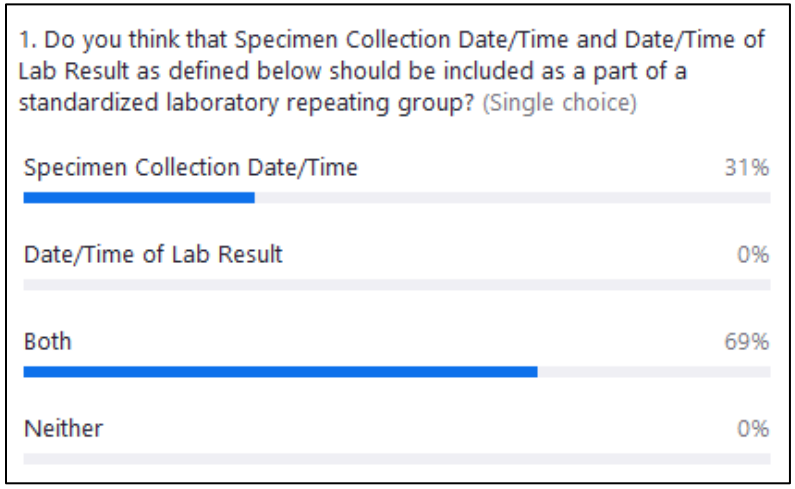
- Is there anything anyone would like to comment on about this data element and its being a part of a standardized laboratory repeating group?

  Discussion
  - Suzanne Gibbons-Burgener asked to clarify that the latest result released for reporting refers to a specific test result. If preliminary results were sent earlier, this latest result would represent the outcome, potentially overriding any previous findings.
    - Erin Kough clarified that, yes, this is the latest information, and it was pulled directly from the brief.
    - Suzanne Gibbons-Burgener replied that It's interesting because we capture the date of result, date resulted, and date reported, which can sometimes differ in our system. It may reflect the date the lab obtained the result, but their system might not trigger the release until the next day or later. I see both dates mentioned in this definition, so that's what we follow.
    - Erin Kough noted that she believed the idea was that it referred to when the result was released—essentially when we used it and that this is a good point. Also, the result date is one thing, but the report date is another, and that was left out of the proposed data elements. We'll take that into consideration, as we want to keep it consistent with the brief.
    - Molly Crockett believed when we initially discussed lab elements, we included the date reported. However, the group wanted it to represent when the result was available to public health, not necessarily when public health received it. This was to account for any delays in reporting, whether due to the reporting mechanism or inter-jurisdictional issues. Essentially, it refers to when the result became verified rather than when the transmission occurred, which can be different. It's a subtle distinction.
    - Tricia Aden commented that the availability of the result might be part of the message's metadata rather than a specific data element. She also commented that the data element and its definition seem to contradict each other. The date and time of the lab result would be in the LIMS when the result is entered and approved, but the definition refers to the reporting date. Does this mean the date it was sent for reporting or the date the lab result was approved and submitted in the system?
    - John Satre was wondering if the interest would be in the clinical date as opposed to a process date for this issue.
    - Suzanne Gibbons-Burgener noted that it's not considered available until it's officially received. She sometimes gets phone calls with preliminary results, but those aren't captured in an ELR—they're not convenient or standardized. We're focusing on what's available through the ELR, which is the date the test result is sent. This can vary depending on whether the lab is the primary performing lab or if they referred it to another lab, as we sometimes receive reports from both. The release date to public health is considered the date it was technically reported, which differs from when it was released within the medical system.
    - John Satre commented that he felt a laboratorian would likely say they released the result because it was verified.
      - Tricia Aden agreed.
    - Carmen Pugh noted that from the laboratory perspective, once a test result is entered into the LIMS, that marks the result date and time. This is distinct from when the result is released to the ordering provider or public health, as those dates and times can vary. The result date and time refer to when the test was completed and recorded, while the report date and time indicate when it was submitted to the ELR for transmission to the receiving entity.

- John Satre replied that his point is that we are interested in the very first date when the lab officially releases the result. Wouldn't that be the relevant date?
- Carmen Pugh responded that she cautioned against using the term 'released.' While a result can be entered into the LIMS, it doesn't mean it has been released to the ordering provider, as release dates and times can vary by facility. Some facilities prefer immediate notification, while others may only want results sent once a day at a specific time. It's important to differentiate between when the result is available in the LIMS and when it is reported to the provider, facility, or public health.
- Erin Kough noted that as she reviewed the brief, it states that the date and time of lab results should reflect when the latest result was released for reporting. So, we do have that language included. (For reference, this is located on page 20 of the Brief, which can be found here.)
- Tricia Aden noted that she preferred the definition that Carmen mentioned for released meaning eligible for reporting as it helps the understanding of this definition because otherwise it is, it can be easily misinterpreted.
- Erin Kough noted this would be taken into consideration.
- Heather Crawford asked to clarify if we're focusing specifically on OBR.22 for this discussion, or are we defining things more generally? She is trying to apply this from my perspective as someone who receives ELRs.
- Molly Crockett noted she didn't have the Brief in front of her but recalled that, at one point, we specified certain locations within the message. However, this was probably the longest and most complex discussion we had about any of the initial data elements, as it became unclear. The focus eventually shifted more toward the meaning, and she couldn't remember if we kept those specifications in the end.
    - Erin Kough responded that she did not see anything and believes the DSWG took a more agnostic approach in the brief and focused on the definition, as Molly suggested we might have done.
- Heather Crawford agreed with Carmen that we need clarification. The information in the HL7 message refers to what happened with the specimen and the results, not when the message was sent to the state.
    - Erin Kough agreed and noted that for the definition, we are tied to the definitions located within the Brief, but this can be clarified within the recommendations document.

Polling

Poll 1 (n= 26) – Specimen Collection Date/Time Date/Time of Lab Result



Proposed Data Element: Performing Laboratory Specimen ID

- Next up is Performing Laboratory Specimen ID, defined as a laboratory-generated number that identifies the specimen related to this test.
- For jurisdictions able to port ELR information into case notifications, this would map to the SPM-2.2 segment.
- If the definition makes sense to everyone on the call, let's go ahead and get started with the discussion. I've listed out some potential discussion points here but am happy to hear from anyone about what they think about having a place in a standardized lab repeating group for this type of data element.

  Discussion
  - John Satre noted that Iowa will provide a value but will not supply the actual performing lab's specimen ID. There are complexities, as the reporting entity is sometimes not the performing lab, and they may use their own accession numbers.
  - Molly Crockett commented that there are various reasons for concern, as it's considered too much of an identifier. Traditionally, we haven't supplied this for CDC reporting. While it might not be a permanent barrier, it would require discussion in Massachusetts before being included in messages.
  - Suzanne Gibbons-Burgener wondered about the expected utility of having this number. Are we already providing a proxy, like the state case ID number, that could serve the same purpose? She can sometimes tell if the same specimen was used for multiple tests based on the collection date and specimen type. We capture the accession number, which is typically provided by whoever reports through ELR, and it can be shared across multiple tests. I'm just curious how many people will use it and what its utility is.

- Erin Kough noted that in some cases, like COVID, there were multiple tests on the same specimen type in a day, though that might be rare enough not to worry about. I'd like to hear from anyone on the CDC side about how the accession number might be used.
    - Peter Boersma replied that he doesn't represent a specific program but has spoken with various programs on this topic. There are more focused use cases for this data element rather than broad application. One example is when testing is done at the CDC—this number allows for a linkage that would otherwise be difficult to establish.
    - Suzanne Gibbons-Burgener commented that this is likely a very specialized use case, so making it a core requirement is probably unnecessary.
    - Erin Kough noted that one thing to consider is whether it makes sense to reserve space for this, even if it's rarely used. It's worth thinking about when to build it and send it.
- Sandra Roush shared that the performing lab specimen ID, along with related data elements, is included in vaccine-preventable disease MMGs (measles, mumps, rubella, pertussis) managed by NCIRD, though not for hepatitis. These elements allow jurisdictions using APHL, CSTE, and CDC-supported VPD lab reference centers to electronically consume lab data directly into case reporting or notification systems without manual entry. While this is a specialized implementation, it was suggested by jurisdiction partners to ensure flexibility for electronic data exchange, even if not widely used.
- Sara Johnston replied that, as Sandy mentioned, this issue relates to linking lab data with case data at the CDC without other identifiers. We've encountered this challenge across multiple conditions, often involving linking NNDSS or case data with data from CDC or public health labs that aren't sent to CDC. We have two options: a more specific approach for identifying labs or a blanket approach. The blanket approach would involve sending the performing lab specimen ID routinely, regardless of the lab, so you wouldn't have to worry about this case by case. While this blanket approach might not be ideal, a more targeted approach would require additional effort. She wanted to hear thoughts on whether a blanket or targeted approach is better for linking lab and case data.
- Tricia Aden noted that additionally, it's not just about linking lab data to case data; it's also about linking specimens submitted to the CDC for reference testing like the VPD reference centers Sandy mentioned. It’s becoming increasingly challenging to do this. The case ID number doesn't come through on the forms sent for CDC specimen testing, so this ID has been useful for internal linking. We lack other effective methods due to the data being mostly de-identified, so she wanted to emphasize that.
    - Suzanne Gibbons-Burgener followed up regarding the last comment about CDC labs not including the state ID number; she understood that this field was required when we submit data. It seems like that ID would be essential for linking internally at the CDC. She was referring to the case ID or disease incident number, which most jurisdictions use for a patient in relation to a specific disease event, not the specimen ID number.
    - Tricia Aden replied that during the monkeypox response, we were unable to use it, which was my most recent experience with data connection. She encountered many challenges, so we would need to examine this more closely.
    - Suzanne Gibbons-Burgener responded that initially, during an outbreak or emerging situation where only the CDC performs testing, we face challenges because we don't receive ELRs from the CDC. Sometimes, approval for testing

occurs before the case information is entered into our electronic systems. There could also be utility in cases where labs performing tests requested by the CDC automatically forward results, especially when looking for variants of pathogens like COVID. In such instances, knowing the specimen or patient ID number used by the laboratory would be valuable.

- Tricia Aden noted that from her experience, the patient ID number is usually not very helpful, but she may not have the full picture. This isn't just about funding new infectious diseases or obtaining approval for testing; it's also about ongoing testing throughout a response. For example, during the monkeypox response, states were asked to send in a specific number of specimens every week or two weeks for follow-up testing, and connecting those pieces is part of the process. It's more complex than it might seem, and we may be discussing similar issues at various stages.
- Sara Johnston replied to support what Tricia said regarding Mpox. She's noticed that the case ID is not consistently included with the specimen form sent to the CDC, making it difficult to link lab and case data. There isn't a simple solution to this issue; if it were easy, it would have been resolved by now, as it's a significant problem. We likely need to improve the process on both sides: ensuring that case data is sent from labs to the CDC with the case ID included in the lab forms while also including the lab ID in the case data to cover more scenarios and accommodate different jurisdictions setups. She emphasized that there is a significant gap at the CDC. The solution should not require jurisdictions to send additional epidemiological content with lab data; rather, it should focus on linking the data already being sent.
- Erin Kough was curious to hear jurisdictional input on the best approach for this issue.

Proposed Data Element: Sequence ID

- This is a more specialized identifier, the Sequencing ID. This is a genomic sequencing ID assigned by institutions such as NCBI Biosample or NCBI SRA upon accession of the sample.
- A large question for this data element would be the generalizability of this data element. Is this data element common enough to be needed in every lab repeating group? Again, this would be a placeholder for use when needed.
- Another question is if the Sequence ID number is more appropriate at the case level rather than the lab level. How common is it that more than one specimen per case is a) sent for WGS sequencing and b) ends up with disparate results?
- Would this be more appropriate as a potential value for the Identifiers data element?

Discussion

- Suzanne Gibbons-Burgener was curious about how frequently we receive the requested sequencing ID number within the timeframe that epidemiologists are working. Do we typically leave cases open, hoping the results will come through? She wondered how other places have incorporated this number into their systems. If we aren't reliably integrating it, she was unsure how we can pass it on effectively.
  - Misty Johnson replied that as an informaticist, she could share that the epidemiologists she has worked with often do not see the sequencing ID number frequently. For example, when specimens are sent for sequencing, sometimes they go to a CDC lab,

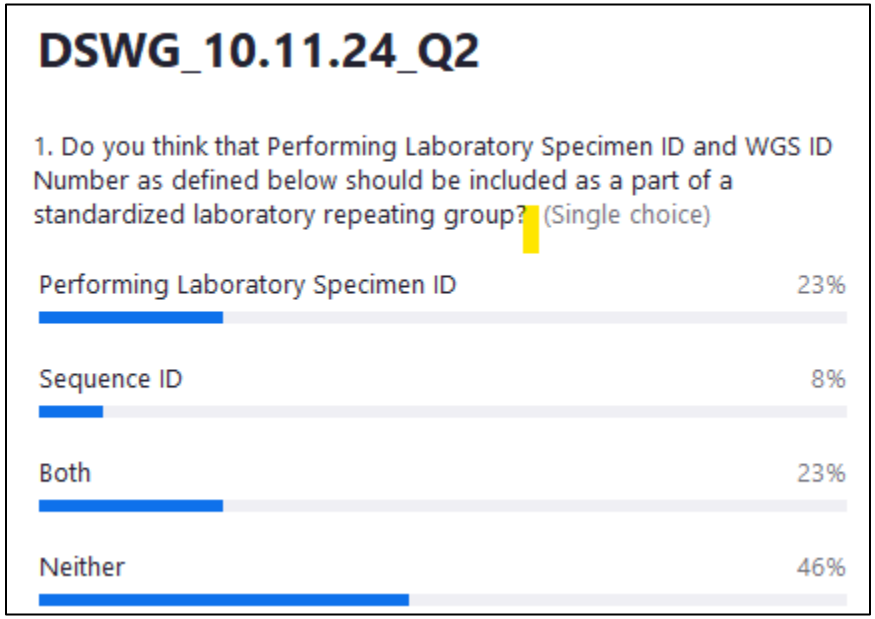
and the results usually come back on paper. As a result, those results are hand-keyed into our disease incident surveillance system. From Wisconsin's perspective, while we may occasionally receive this information, it's relatively rare based on my observations. It's certainly not common, especially not from the ELR.

- Suzanne Gibbons-Burgener noted agreement that she was thinking about instances where we're advancing sequencing programs, such as with wastewater, which differs since we're not dealing with individual patients. In outbreak situations, we do get sequencing to check for matches, and the CDC uses this approach as well, particularly with foodborne outbreaks. She sees the utility in it but was unsure if we're fully prepared to implement it yet. We may not have caught up with how to make this happen, but we recognized its value, especially at the national level, for tracking multi-jurisdictional spread or the magnitude of specific strains.
- Tricia Aden commented that she believes we are building it for a future state but also developing it gradually. As more processes move into the sequencing realm, its relevance and prevalence will increase over time.
- Suzanne Gibbons-Burgener replied that in addition to the sequencing ID, we may need to specify the types of entities involved. For instance, do all these entities have specific identifiers like "SAMN" and "SRX" that inform the data recipient of their origin?
- Sara Johnston noted that as we transition to incorporating more sequencing data, as Tricia pointed out, its importance will increase. The question is whether it makes sense to include the sequencing ID in the lab repeating group. Do we have the capability to send it when needed, or would that be too challenging? Currently, there are no lab observations tied to the sequencing ID, only the specimen ID. Should we consider the sequencing ID more like a general identifier, as discussed earlier within the DSWG, rather than linking it to lab observations? This would involve treating it as a genomic sequence ID and specifying which type of ID we are referring to. If the sequencing ID is not currently part of your observations, but we're asking you to send it, how can you transmit it? Would it be feasible to include it in a lab repeating group without any other related content? Alternatively, would it be more effective to send it as a standalone identifier for now?
- Erin Kough noted agreement and that many systems may struggle to capture the information if it is set up as part of a laboratory repeating group, especially since the data isn't currently being submitted in that format. This is an important point to consider.
- Sara Johnston commented that if we expect to have sequence data flowing through ELR and lab observations with sequence IDs integrated into our systems within a couple of years, we may need to consider an alternative solution for specifying the ID for that data. She realizes this might not align with everyone's perspective, but it's worth exploring different approaches.
- Suzanne Gibbons-Burgener had a question about the sequence ID. When you receive the ID number, is it automatically associated with the specific organism, like E. coli, or does it require additional information typically transmitted in an ELR? For example, does one need details about the origin of the specimen, the date, and other relevant data linked to the test result? Ideally, this would be entered as a separate test result and may not necessarily be linked to the culture, as it could have a different ID number and be processed at a different location.
- Sara Johnston replied that, ideally, it makes sense to have the sequence ID structured that way. However, if jurisdictions aren't currently equipped for this or are progressing slowly, it would be helpful to hear about their advancements. If these aren't

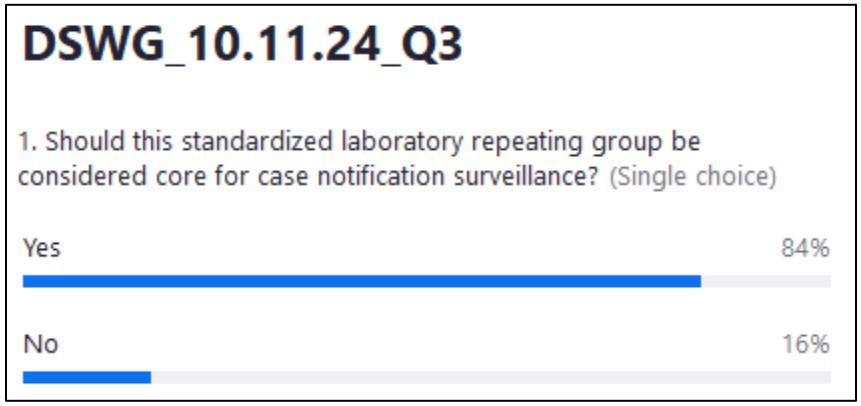
captured as discrete observations in a lab system, the next best option would be to use the ID to link to another system that contains essential metadata, such as dates and additional information about the sequence.

- Tricia Aden replied that the key question is that, when looking at an ID in NCBI, it provides the actual sequence but includes minimal metadata. While she understood the point from Sarah, she wanted to clarify that this sequencing ID number primarily serves that purpose.
- Sara Johnston commented that we also conduct sequencing in public health that is not sent to NCBI or any other repository; it is retained solely within public health systems. These have different specimen IDs, which also adds complexity. Then, these are not linked appropriately, which is the challenge. It's a complex issue, and we won't resolve it in a few minutes. My question for the group is: what steps can we take right now to make progress on this matter, even if we can't find a complete solution?
- Erin Kough moved the conversation to the second poll to help move this issue forward and gain a pulse from the group on where this stands.

Poll #2 (n=13) – Performing Laboratory Specimen ID/ WGS ID Number



Poll #3 (n=19)– Standard Lab Group = Core?



Next Steps

- This concludes the Laboratory DEST discussions.
- An after-meeting input form will be shared with this DEST to gain additional information before the draft recommendations document is shared with the group for review and input.
- JMC will work closely with Co-Chairs and CSTE to draft the document, which will be sent to the DEST for offline review.
- The second Social Determinants of Health call will be held on October 18.
- 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.