Meeting Title: CSTE Data Standardization Workgroup Meeting
Date: 2/28/2025

Attendees: 83

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

CSTE Data Standardization Workgroup Meeting (Co-Chair, Laura Erhart)

Core Data: In Scope for DSWG

- For these meetings, we're most specifically identifying the core surveillance data elements we will look for transmission to the CDC.
- We are looking at the core overall, focusing on case surveillance, situational awareness, or routine surveillance but going across many conditions.
- The data elements we discuss don't have to be relevant for all conditions but should also not be something that's just in one or two specific places.
- We're looking at case-based surveillance again, so it's not syndromic, not aggregate reporting.
- As we go, we'll examine specific selected data classes. Again, the scope identifies common definitions, value sets, code sets, structure, and some interpretation notes.
- For the core data elements and where applicable, we will recommend some best practices on what we may need to collect within our surveillance systems.

FAQs

- Anything we identify for core data for CDC transmission doesn't mean it's required.
- Even if we do identify an element for transmission, it doesn't need to be filled in for every single element or every single case; it's okay to leave some of the fields blank or unknown if the information is not available, not a priority, or not useful for the condition.

Considerations

- What information is needed at the national level?
- How can we use electronic records and automation in our current electronic digital age and landscape?
- Also, how can we best take advantage of and feed into other standardization efforts like HL7, USCDI, and USCDI+? None of what we do in public health happens in a vacuum.
- How can we reduce the manual surveillance burden on ourselves and our partners?
- Finally, how do we position ourselves well for adding emergent conditions, exchanging information, onboarding, case notifications, etc.?

Meeting Ground Rules

- We welcome active participation and engagement from everyone on the call.

Announcements and Key Updates (CSTE, Meredith Eddy)

Announcements and Key Updates

- Reminder: CSTE Brief – is open for comment through 3/12/25.
- Coming soon to CSTE Connect for Level 3 review: Travel Core Surveillance Data Class – eta mid-March

Status Updates for Recommendations (JMC)

Status update for closed DESTS

- Ten of the data classes have had their recommendation documents handed off to CSTE for the next steps in the longer review and approval process. Travel is the first document to go to the level 2 review by the steering committee and subcommittee chairs, and JMC and CSTE are working with the co-chairs to address comments before it's released for Level 3 review by general CSTE membership.
- The document for Signs and Symptoms is out for review by the DSWG, and that feedback is due March 7th.
- The document for Laboratory is being cleaned up before it's finalized for handoff for Level 2.
- Exposures, Risk Factors, and Underlying Conditions is dependent on Signs and Symptoms to ensure alignment since some of the conversations around structure were intertwined.

Recommendations Summary (as of 2/28/2025)

- This is a visual summary of the status presented on the previous slide.
- Incredible progress was noted for this groups efforts!

SDOH: Unknown (JMC – Sam Spoto)

SDOH Data Elements

- Here is an overview of the five data elements the DSWG has decided are core data for case notifications to CDC.
- This includes Housing Status, Living Situation, Preferred Language, Educational Attainment, and Health Insurance or Access Type.
- An item that came up routinely during our discussions of each of these was what to transmit as a value when the answer is not known.

CDC Methods for Conveying Unknown/Missing

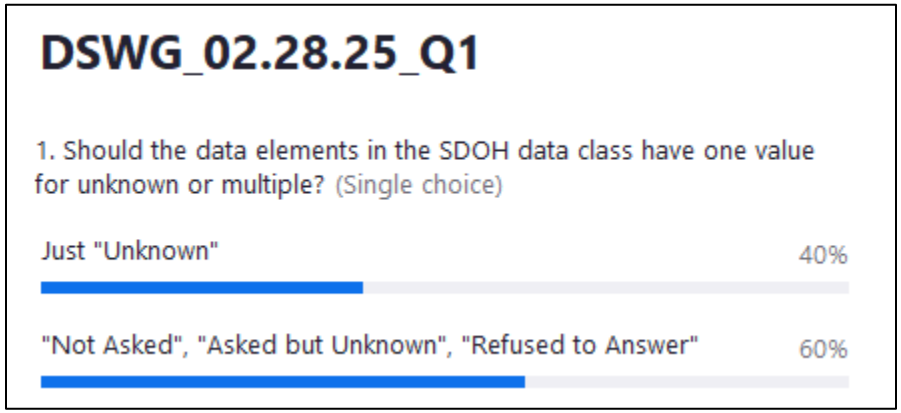
- In the data classes we've tackled over the last year and a half, almost every data element only uses one value to express "Unknown". The two exceptions here are in the SOGI data class, where some data elements have an option for "choose not to disclose".
- There is existing CDC guidance on Methods for conveying unknown or missing values in case notification messages. This guide focuses on the distinction between unknown and blank, which is a little different, but the slide contains a few sentences here where the sentiment they express could still apply.

    - The guidance talks about the burden placed on jurisdictions to modify their systems if required to collect information on unknown types, as well as the additional burden in data collection.
    - The final sentence reads "The NNDSS Program recommends that CDC Programs be judicious in requesting that submitters distinguish between “unknown” and “missing” values in HL7 messages so as to minimize the burden on the jurisdictions."
- We want to make sure we cover this guidance to provide some context for how things have been set up currently.

Unknown Discussion

- As a reminder, whatever is decided here today applies only to the information transmitted for these core data elements to the CDC.
- Given all that information, we want to hear thoughts from the group on the public health use case for conveying not asked, asked but unknown, or refused to answer rather than just one option for unknown for these five SDOH data elements.
- Do we think everyone would fill out these choices the same way, or would it take additional guidance to know everyone is doing it the same way? Is adding this complexity taking away from the goal of standardized data?

Poll 1 (n=43) – SDOH Unknowns



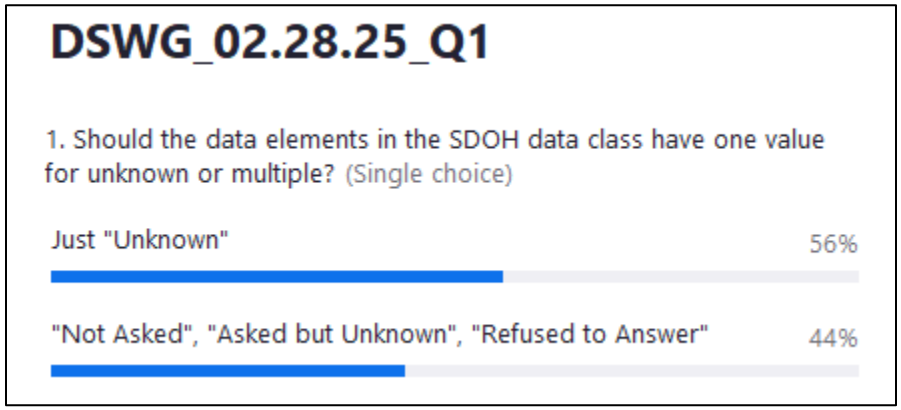
Discussion

A discussion ensued and focused on the granularity of "unknown" values in data collection. While most data elements use a single "unknown" value, some, such as SOGI, have multiple distinctions (e.g., "chose not to disclose"). Participants were divided on whether more granularity is beneficial, but many leaned toward including it for clarity. One example from a jurisdiction’s drug overdose reporting group showed improved data quality after adopting a more detailed "unknown" classification. There was concern that states might be judged based on the proportion of unknown values, making it important to differentiate between cases where data was not collected versus cases where individuals refused to provide it. A key question raised was whether this approach should be standardized across all data elements or remain specific to certain fields. It was acknowledged that jurisdictions could collect more detailed data while still reporting a simplified version to national databases. However, without national representatives on the call, the long-term implications for national reporting remain uncertain.

The discussion concluded by acknowledging that while granular data can be mapped to a general "unknown" category, the reverse is not possible, making it important to decide on standardization carefully. The discussion highlights that while some local jurisdictions may find value in collecting more granular data, submitting data at a less detailed level seems sufficient overall. Standardizing this across all jurisdictions would be a significant effort, and in some cases, certain data points would remain unknown despite enhanced collection efforts.

After the discussion, the poll was launched again to gather the group's thoughts.

Second Launch **Poll 1 (n=43)– SDOH Unknowns (n=43)**



SDOH: Living Situation

SDOH Data Elements

- Today the other goal is to finalize living situation. There has been a lot of conversation about the value set, and people have a natural question when thinking about at what point a place, like a hotel, is considered someone's living situation. To answer this question, we need to consider the purpose of this data element.
- By defining a data class for SDOH, we are trying to do something different than what we've historically done in case investigations. These data elements add (not replace) information in case notifications that we think can allow for transmission of data that promote health equity work.
- We've said before that these data elements should not aim to capture an exposure or risk factor for a condition, as those concepts can be transmitted in exposures or risk factors.
- Similarly, our SDOH work shouldn't focus on capturing where someone may have spread an illness. The analysis of SDOH data elements should focus on whether a certain population is disproportionately affected by a condition.

SDOH: Living Situation

- Here is our current definition and value set for living situation.
- In future slides, we'll focus on specific sections.

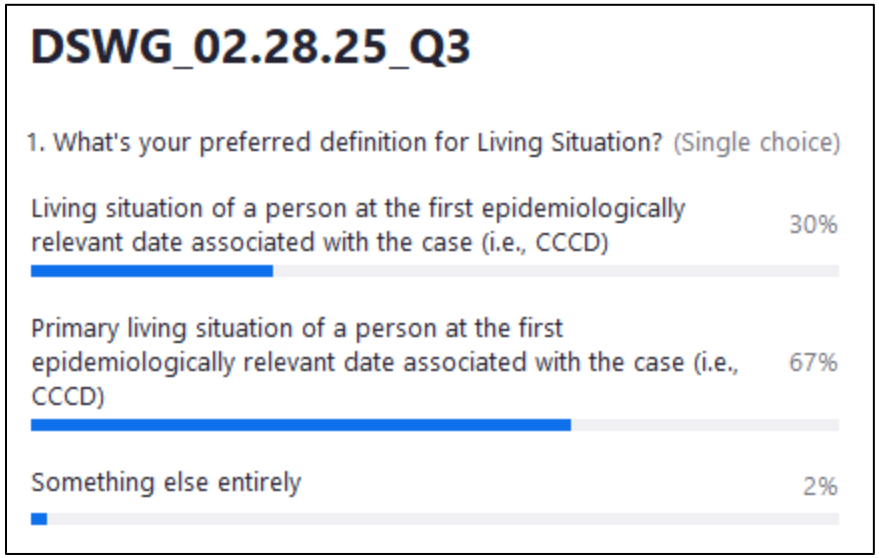
SDOH: Living Situation (continued)

- First, we'll discuss whether our definition makes enough sense that everyone will fill it out in the same way, and then we'll discuss some very specific value set feedback we received.
- It may be unclear when something should be selected as a living situation, such as an acute healthcare stay, or someone on vacation in a hotel, or in jail for a day or two. One method of "standardization" here could be defining scenarios about what should be selected and when. Still, we need to consider how this might be implemented in a future MMG and if we think case investigators across all conditions and jurisdictions would both read and follow that level of detailed guidance for this data element.
- During the feedback process, one idea was to specify that we are asking for someone's "primary living situation". As part of this update, we'd likely remove the option for an inpatient hospital or acute care facility as it's extremely unlikely that that would ever be someone's primary living situation. It could fall into our "other" category if it did exist.
- Do we think calling this a "primary living situation" would help others understand what kind of data should be transmitted here? Is that enough clarity to ensure people will answer this consistently?

Discussion

The discussion focused on clarifying the definition of "primary living situation" in surveillance data collection. Several participants supported "primary" in distinguishing stable living arrangements from temporary situations. A participant raised concerns about potential overlap with an existing "setting" variable, which already captures locations like healthcare facilities, schools, and jails. However, the distinction was made that the "primary living situation" focuses on where a person resides rather than where they spend time for other purposes. There was debate about defining living situation at a specific epidemiologically relevant date or as a more stable concept. Some pointed out that certain cases, like student housing or long-term care, fall into a gray area between temporary and permanent. Others noted the importance of tracking movement, as people may frequently change residences. The conversation also touched on short-term housing arrangements (e.g., couch surfing) and whether stays of 14 days or less should be considered "primary." The group acknowledged the complexity of defining and applying these concepts consistently, with a plan to further refine specific terms and values.

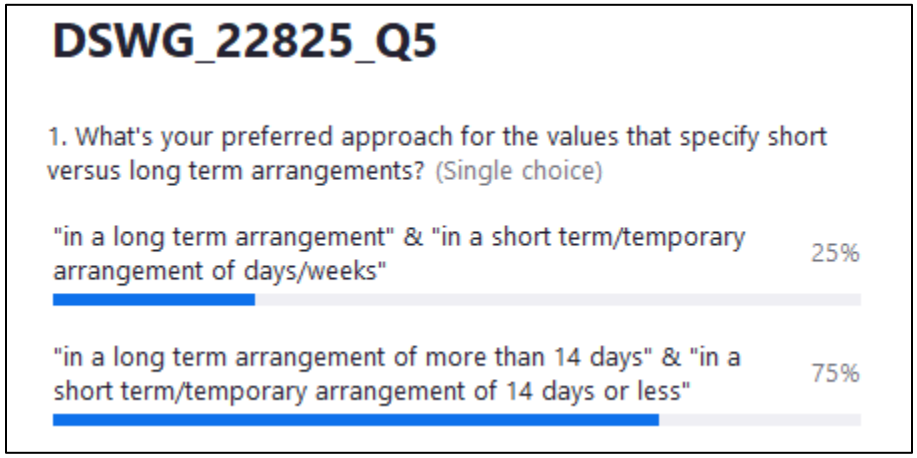
Poll 3 (n= 43) – Living Situation Definition



SDOH: Living Situation

- Now, we'll look at the value set. In Living Situation, there are some responses that specify a long-term arrangement and some that specify a "short-term arrangement of 14 days or less". We received some competing feedback here. There was some concern that people would get too focused on figuring out an exact length of stay if we enumerate the values like "14 days" or maybe not know what to choose if you don't know for sure the exact duration of the living situation, versus others who said we specified 14 days or less on short term. Hence, we need to specify 14 days or more on long term. This is a great example of one of the most common dilemmas we face in the DSWG. For standardization purposes, we want to make definitions that everyone naturally implements in the same way or as close as we can get to the same way. Still, we don't want to make things so precise that we have to select "unknown" when we might know an answer that's probably good enough. For example, you might read in a medical record that someone is couch surfing with friends, but have they been in their current situation for more or less than 14 days?
- Does anyone have any thoughts here on the most implementable solution, thinking of both standardization in data collection and interpretation for analysis?

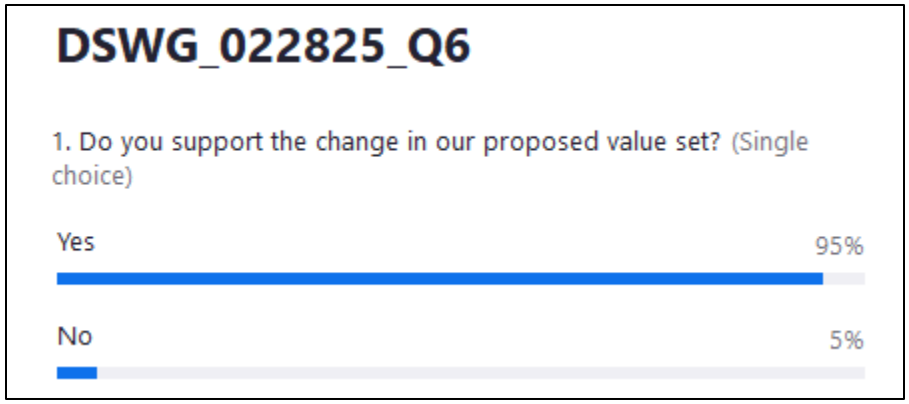
Poll 5 (n=32) – Short/Long Term



SDOH: Living Situation

- We also received feedback about the values for some unstable housing arrangements. Part of the feedback was a need to better separate what might be a congregate or non-congregate setting, and other feedback was that having too much granularity in these options would be difficult for data collection. The proposed changes here are to split out a shelter or haven that's a congregate setting from other temporary non-congregate housing provided by charity or government programs, like transitional housing or pallet shelters. Additionally, the options for structure not meant for human habitation and vehicle not meant for human habitation have been combined, and the two options for open air have been combined. We also added a value for "vehicle meant for human habitation", like an RV, since a few people had questions about where that would exist.
- Do these changes balance the public health needs here?

Poll 6 (n=37) – Unstable Housing Options



DSWG Next Steps: (JMC)

Next Steps

- The next meeting is scheduled for March 28th, and the topic has yet to be determined.
- Reminder that if you'd like to review the draft for signs and symptoms, it has been sent out to the group, and feedback is due next Friday, March 7.
- Soon, participants will see the SDOH draft for DSWG review.