Meeting: CSTE Data Standardization –Social Determinants of Health-Data Element-Specific Team (DEST) Meeting
Date: September 20, 2024
Attendees: 97

CSTE (3)	PHAs (55)	CDC (33)
Becky Lampkins	DSWG Co-Chairs	Akaki Lekiachvili
Megan Sullivan Toe	Judie Hyun | MD	Amy Board
Meredith Eddy	Laura Erhart | AZ	Ashley Meehan
JMC (2)		Austin Brown
Laura Lane	DEST Members	Chirine Chehab
Sam Spoto	Allison La Pointe | MN	Danielle Sharpe
	Andria Apostolou | IHS	Danielle Tack
Unknown (4)	Ariana Beattie | MN	Henrietta Smith
Charlene Adams	Ben Meana | WA	Jeanne Ocampo
Danielle Boothe	Brianna Rossi | VA	Jennifer Wiltz
Katie McNabb	Carlene Robinson | AL	Jessica Purser
Yenling Ho	Chelsea Robinson | WI	John Gerstle
	Cher Levenson | WA	Josh Bagley | GDIT
	Deb Loniewski | MI	Kristie Clarke
	Devann Kirkpatrick | TN	Lauren Korhonen
	Ed Hartwick | MI	Mariel Marlow
	EJ Klingler | NYC	Marion Rice
	Elana Silver | CA	Meeyoung Park
	Elizabeth Kessler | NV	Myrna del Mar González | GDIT
	Elizabeth Schiffman | MN	Noreen Kloc
	Eman Addish | Philadelphia	Peter Boersma
	Erin Doyle | NG	Phoebe Thorpe
	Genie Tang | CA	Rachelle Oseni
	Hanna Chakoian | MN	Rahsaan Overton
	Hollie Sheller | MO	Renee Boehmert
	Isaiah Francis | NM	Robert Pratt
	Jacob Chavez | LA	Sandy Price
	Janet O'Connor | SC	Sara Johnston
	Jason Singson | CA	Seung Hee Lee
	Jennifer Kwon | LA County	Sharon Silver
	Jesse Ellis | MS	Shawn Arshad | GDIT
	Josh Leopold | MD	Sophia Kazakova
	Kayleigh Nerhood | GA	Walter Alarcon

	Keely Wrolstad | WI Leah Porter | MN Li Jiao | Harris County Linda Kelemen | SC Lunden Espinosa| IN Mario Beltran | MDHHS Maya Scullin | OH Elizabeth Lando-King | MN Melissa Kealey | CA Miriam Van Dyke Misty Johnson | WI Orlando Velazco | CT Pyone Yadanar Paing | WA Riley Gulbronson | TN Robbie Madera | Philadelphia Sarah Love | Chicago Shannon Allain | OR Shawna Stuck | GA Steve "Ike" Eichner | TX Suzanne Gibbons-Burgener | WI Syreen Goulmamine | VA Theron Huntamer | NV Tia Bastian | MN Trish McLendon | CA Tyler Riedesel | UT	

**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***):**

Social Determinants of Health

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.

Core Data FAQ

- When voting for core data, consider that naming something as core does NOT mean that it will be required to be populated for every case to CDC.
- It is okay to leave a field blank or unknown if the information is not available or not applicable to certain conditions.
- A data element should be considered core if it's relevant for many conditions, although there is no set number for "many."

Social Determinants of Health

SDOH versus Health-Related Social Needs

- In public health frameworks, a distinction has been made between SDOH at the community level and health-related social needs at an individual level.
- This distinction is acknowledged here and clarifies that we are focusing on person-level data because case data are collected at the individual level.
    - While our naming convention for what we're talking about has been SDOH, what we are attempting to capture with these data elements are individual health-related social needs.
- These data points help public health officials identify which individuals may be at higher risk of adverse health outcomes due to their social conditions but also highlight factors that may increase the risk of acquiring and spreading disease, particularly in congregate living settings**.**

Prioritization of Concepts for Discussion

- Health-related social needs are a broad topic that encompasses a wide range of data elements. To gather information ahead of the call, a poll was sent to the DEST members on the individual-level concepts for which it is most important to define standard definitions and interpretations for case surveillance.
    - The poll identified homelessness/residence type and incarceration as some of the topics of highest priority for discussion. For the first call, we'll focus specifically on residence type as a data element for how homelessness and current incarceration or detention may be captured and transmitted in case surveillance.
- Future calls will cover other high-interest areas, including:
    - Primary Language
    - Highest Level of Education
    - Insurance Coverage and Income Level
- Topics that will not be covered because they have already been discussed or will be discussed in other data classes include industry and occupation (I&O), SOGI, Exposures/Risk Factors such as substance abuse, Disability, and Race and Ethnicity.

Background Information: SDOH – Residence Type

In this presentation, we will take a closer look at how various standards or workgroups have defined data elements related to Residence Type, particularly homelessness and residence in a congregate setting, including correctional facilities.

Concepts Related to Residence

- The conversation will focus on transmitting key aspects of residence type to understand the health risks associated with different living conditions in the context of reportable diseases.
- When we talk about the data element of Residence Type, the two concepts that are in scope for transmission by that data element include:
    - Homelessness: Includes individuals without a stable or permanent home. Homelessness can be divided into two categories- sheltered or unsheltered.
    - Residence in a Congregate Setting:
        - Correctional/Detention Facility: Living in jails, prisons, or detention centers, where close quarters can heighten the risk of infectious diseases.
        - Other Congregate Settings, Such as residential care facilities or group homes, where shared living spaces present unique public health challenges.
- Concepts related to residence that are out of scope for the data element we're defining today include:
    - Housing Stability/Instability
    - Housing Quality/Conditions
    - Crowding
- By narrowing our focus, we aim to collect a standard data element across cases to better understand how particular residence types, like homelessness and congregate settings, impact reportable diseases.

The Gravity Project

- The Gravity Project is a multisector initiative dedicated to standardizing Social Determinants of Health data, and one of its key areas of focus is housing/residence. This includes developing standardized data elements for housing instability, homelessness, and overall housing status.
- The Gravity Project designs data elements to be integrated with Electronic Health Records (EHRs) with the goal of eventual incorporation into the United States Core Data for Interoperability (USCDI). The Gravity Project's work is also closely aligned with CMS initiatives like those we just mentioned.
- The Gravity Project defines and categorizes homelessness to ensure accurate data capture and reporting.
    - Sheltered Homelessness: Includes individuals currently living in shelters, motels, temporary or transitional housing situations, or scattered-site housing.
    - Unsheltered Homelessness: Refers to individuals residing in places not meant for human habitation, such as cars, parks, sidewalks, abandoned buildings, or on the street.
- The Gravity Project is also planning to expand housing work this year by developing standards and terminology around a CDC-developed and validated question set to capture residence data, focused specifically on homelessness and residence in congregate settings.

CDC- Validated Question Set

- CDC and Fulton County, Ga have worked together to develop and validate a question set for accurate capture of housing status (whether someone is housed, experiencing sheltered homelessness, or experiencing unsheltered homelessness) or living in a congregate setting.
- An idea of the flow of questions and responses was shown on the slide. The person should be asked to provide information about where they are living in their own words, and the investigator then would select the appropriate options from an extensive list. Depending on those selections, additional follow up questions may be required, such as is the location made for a human to live in or paid for by charity.

CDC- Validated Question Set (continued)

- The responses from the previous questions can be rolled up into the 17 options shown on this slide.
- The Gravity Project is working to develop standards and terminology around this set of values for future inclusion within USCDI and electronic health records.

ELR & eCR

- In electronic reporting to public health, both ELR and eCR contain a patient address. If the address listed is that of a known homeless shelter, this can suggest that the individual is experiencing housing instability, or if it is of a correctional facility, you may be able to tell that someone is likely incarcerated. In some cases, the address field might include a term like 'homeless.' In other instances, it may contain terms like 'no fixed address' or a general delivery address at a post office, which are not necessarily but can be indicators that the individual does not have a permanent residence.
- In eCR, 'Characteristics of Home Environment' can contain data on the type of residence, the status of accommodations, living situation, and environments. For example, the value set contains SNOMED codes like 'Homeless family (finding)' or 'Housing instability (finding).' The value set for this data element consists of 278 detailed SNOMED values, and not all of them are applicable to our discussion of residence type.

Case Notifications to CDC: Message Mapping Guides (MMGs)

- Specific CDC Message Mapping Guides (MMGs) include data elements on Residence and Homelessness. For example, the Foodborne and Diarrheal Disease MMG describes homelessness as having no fixed residence for any given period, while the Respiratory and Invasive Bacterial Disease (RIBD) MMG asks, "Was the patient homeless at the time of symptom onset?"
- When speaking about Residence as a data element, the RIBD MMG asks "Where was the patient a resident at time of initial culture? “
- Another MMG that includes Residence as a data element is the COVID MMG but instead of describing where was the patient a resident at time of initial culture, it describes where the patient was staying at time of illness onset.
- Not shown in this slide is TB. In the TB MMG, the risk factor data element includes values for “Homeless Ever” and “Homeless in the Past 12 Months.”

Case Notifications to CDC: Message Mapping Guides (MMGs) (continued)

- The Hepatitis MMG also asks about the type of incarceration facility where the patient was incarcerated for longer than 24 hours before symptom onset. The TB and Latent TB MMG includes the Type of Correctional Facility data element which describes the type of correctional facility if the patient was a resident at correctional facility at the time of diagnostic evaluation.

Case Notifications to CDC: Message Mapping Guides (MMGs) (continued)

- The Hepatitis MMG includes various data elements all related to incarceration. The data elements describe the period an individual has been incarcerated or if ever incarcerated.

- The STD MMG also includes a data element related to incarceration: Been incarcerated within the past 12 months.
- Many of these specific examples are of data elements that tend towards condition-specific timeframes and risk factors. The data element we are going to talk about today of "Residence Type" is not intended to encompass or replace all condition-specific risk factors or exposure data elements that would be transmitted within those condition-specific repeating groups.

Current Residence Type

- To that end, the concept captured here is the CURRENT residence type (or perhaps the residence type at the time of the report, time of dx, time of CCCD, or similar permutation). This may be different than what the residence type was at the time of exposure or what a prior residence type was as a risk factor for the condition, as these exposures or risks can be captured within the condition-specific repeating groups as needed.
- As stated earlier, the goals for the residence type are to transmit information on someone's housing status and whether they live in a congregate setting, including incarceration.

CCCD

- The CCCD is defined as the earliest available of six component dates associated with a case. These dates are listed on the right side of the slide and encompass a range of dates between symptom onset and the date the case is entered or recorded. The important feature of CCCD is that it is possible to calculate a value for it for all cases, even if all components aren't available for each case.

Discussion: SDOH- Residence Type

Proposed Definition

- The definition we're starting with today is for Current Residence Type, defined as the "Current living situation or housing status of a person at the time of the CCCD." This would be a coded data element with a value set that we'll determine on this call. Before we start the discussion, we'd like to share two potential value sets so everyone is on the same page with what this data element could capture.

Residence Type Value Set (Future Standard)

- One potential value set listed here relies on the work being done by the Gravity Project that we mentioned before. The future goals are for these concepts on this slide to have associated LOINC or SNOMED codes and be captured in EHRs through inclusion in USCDI. However, in our current state, this specific level of detail is not captured in EHRs.
- This doesn’t currently align with anything in the Arizona system, for example, but we’ve proposed it as a potential value set that could become our future standard.

Condensed Value Set

- An option for a more consolidated value set is shown here. This is more along the lines of some of the value sets currently used for residence type in condition-specific MMGs. It aims to capture information about housing status and congregate settings that can be actionable while also sticking with a smaller number of options that may be more likely to be ascertained during medical record review in case an interview can’t be conducted.

Discussion

- Started with these prompts:
  - Is this definition of Current Residence Type appropriate?
  - Should Current Residence Type be a core data element?
  - Is the more detailed "Future Standard" or "condensed" value set preferred?

Discussion

- Suzanne Gibbons-Burgener commented that having worked on many of these cases, she understands the concept of the CCCD and what goes into it. However, most people asking questions or interviewing disease investigators may not fully grasp what feeds into it. There can also be variability depending on the type of disease, such as the incubation period or the time it takes to investigate. For example, someone could be in the hospital for an extended period, so the relevant date may not be universally known at the time of the interview. Often, you're still in the process of collecting information, and the onset date of the illness may not yet be available. This can make it tricky for the person gathering information from the patient or medical records to identify the exact CCCD.
- Syreen Goulmamine noted that regarding the timing and what is considered full-time versus part-time, they had used dormitories as an example, along with other situations where an individual may be housed differently. With that in mind, are we setting any time limits? It was understood this could be restrictive in some cases, but is it worth establishing a timeframe?
  - Laura Erhart replied that for the timing component, thinking back to some of the MMG questions that Judie raised, many of them considered timing in relation to risk factors, often with the incubation period in mind. However, this specific case is more focused on current circumstances, as Suzanne also pointed out. The idea is to capture information about what's happening at the time of diagnosis, exposure, or report to public health rather than looking back to the incubation or exposure period. In terms of CCCD, it's important to consider what it means to those conducting interviews and whether that is evolving. There are two points to note here: Like many data elements we've discussed in these sessions, this doesn't need to be the exact question asked during the interview. Instead, it’s more about how we define what is being sent. However, you do need to clarify which aspects are being defined. In the industry and occupation data element set, we had similar discussions about what 'current' means. Ultimately, the decision was to define it simply as 'current,' which might apply here as well, focusing on the time of the CCCD. The challenge, though, is how to define 'current' in a way that's both clear and useful.
  - Orlando Velazco agreed with these comments. They were more concerned about how the question was phrased rather than the definition. For simplicity’s sake, what’s the difference between 'living situation' and 'housing status'? How are they defined, and how do they differ? Also, how does that distinction affect the data element overall? The reason they asked is that they were more concerned about how the questions were phrased. From an understanding standpoint, some people may not know what CCCD stands for or may be unfamiliar with certain terms. So, simplifying any definitions where possible would be helpful.
- Andria Apostolou wondered if an exposure from 3 months ago is relevant to outbreak control or investigating an outbreak at a shelter, then having the history makes sense. So, they were not entirely clear—what is the intent? Is it for the individual to assess their own risk or determine if they received appropriate follow-up treatment? Or is the intent broader, like for outbreak control or other purposes?

For example, if someone is currently living at home but lived in a dormitory two months ago where there was an outbreak, it seems relevant to know where they were 2-3 months ago, especially if it ties into the incubation period.

- Laura Erhart replied yes and no. This ties back to the timing question—what is the timeframe we're considering when collecting this information? Is the type of residence a risk factor for what we're assessing, or is it more about outbreak control and investigation? There are different time periods relevant for various diseases and purposes. In this context, we're discussing a very broad topic under the umbrella of social determinants of health, which can also apply to other uses. There will continue to be options for condition-specific questions and data elements to address the unique needs of different conditions and their use cases. For today's discussion, we should consider whether the residence type question should be included across different conditions. This raises the question of whether it should be a core element and how we define that. There can certainly be additional uses and needs for specific questions.
- Akaki Lekiachvili noted that they would reference the approach taken by the United States Core Data for Interoperability (USCDI) regarding pre-coordinated and post-coordinated data elements. Pre-coordinated data elements define the concept in a single term, while post-coordinated elements separately define the components—such as admission time—allowing them to be used together. There is a preference for using post-coordination when discussing standards. In a similar vein, they wondered if it would be helpful to define 'residence type' separately, along with associated value sets. This way, it can be applied in different contexts, whether it's current within three months or for other programmatic needs. It seems that we are trying to combine two definitions into one here.
- Laura Erhart noted that for today's discussion, we'll focus on what should be considered core. However, as different conditions arise, additional elements may also be needed. She would encourage using the same value sets for those if that works for everyone.
- Ashley Meehan commented that they wanted to revisit Orlando's question and some of the discussion in the chat. They believed that 'living situation' would also encompass 'housing status.' This distinction can help determine if someone meets the definition of being unhoused or homeless. They wanted to note that this could have implications for how the value sets are categorized. For instance, separating 'homeless' from 'congregate setting' might not be the most logical approach. Therefore, keeping it as a 'living situation' would better allow for the assessment of housing status.
- Tyler Riedesel replied that we need to clarify what 'current residence type' really means. They considered various scenarios, such as someone who was traveling and got sick but returned to their home when they began interacting with the healthcare system or public health. So, what does 'current' truly refer to? They suggested defining it as the time when they are interacting with the healthcare system or public health. This includes when they first made contact or started feeling ill. All of this should be captured in their history, particularly for contact tracing and follow-up. So, 'current' should refer to the period of interaction with the healthcare system or public health, as suggested.

- Orlando Velazco agreed that as a core question, this is valuable. Depending on the need for disease investigation, we might be looking for information that happened in the past, which can be captured through additional questions. A core question, like 'as of this date,' is essential because I assume there will be a date associated with any time this question is asked, allowing for accurate coding. As a core

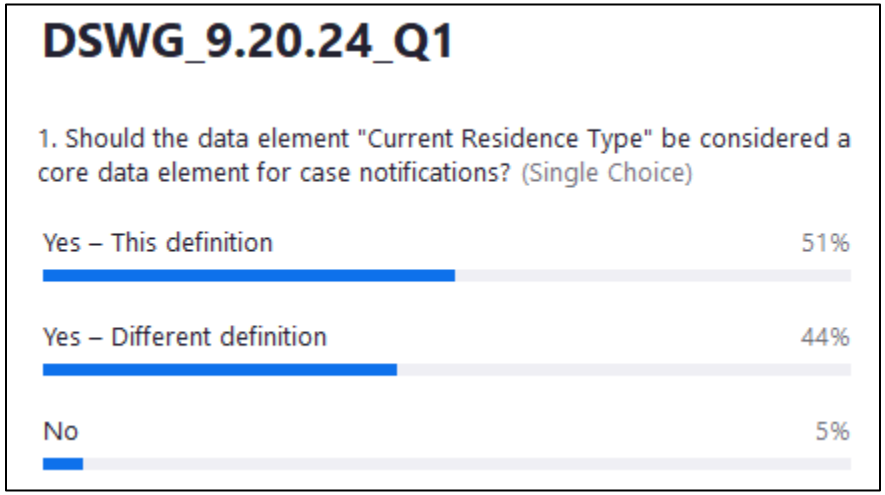
element, it provides important context. They preferred a more condensed approach but believed it was good for inclusion. For example, if someone indicates they're in a shelter during a disease investigation, you will obviously want to find out which shelter. That information can be gathered through further investigation tools. As a core question, this is critical for understanding their current living situation. From a methodological perspective, they focused on the value of this question. At a specific time and date, understanding a person's current living situation is essential because it reflects their current needs. If they report being in a shelter, that might trigger a referral for housing assistance. In summary, this was seen as an important core element that should be included. While the preference was a more concise version, it was understood why some might want more detail. Those specifics can be addressed through follow-up questions.

- Laura Erhart agreed that this question isn't meant to capture everything needed in one go. Depending on the responses and the situation, there will likely be follow-up questions to gather more specific information about locations, duration, and so on. However, we won't delve into all those follow-up questions as part of the core discussion.

Polling

Poll 1 (n= 43) – Residence Type

The first poll was on whether residence type should be a core data element for transmission to the CDC. If you think some kind of residence-type data element should be core across conditions, select one of the "Yes" options to help us see who supports this definition versus a different definition. If you don't think it should be core at all, then select No.



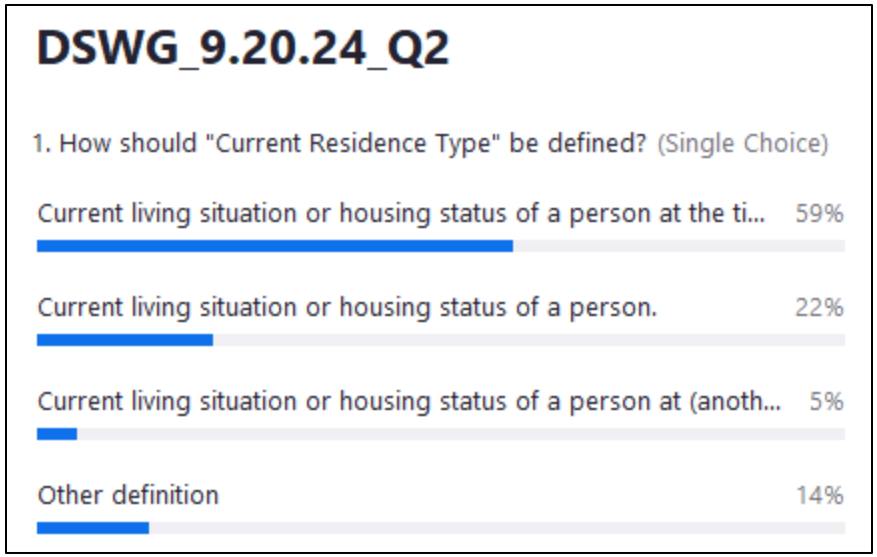
Poll 2 (n= 37) – Residence Type Definition

Polling questions included options for

- Current living situation or housing status of a person at the time of the CCCD.
- Current living situation or housing status of a person.
- Current living situation or housing status of a person at (another date).
- Other definition

Discussion

- Suzanne Gibbons-Burgener commented that instead of using 'CCCD,' we should focus on identifying the beginning of the disease occurrence. If they were speaking with someone, that's how they would phrase it. If the intent is to determine when the disease started, which is the purpose of the CCCD, we need to clarify that. For example, if they were following up with someone who has just been discharged from the hospital, they would ask, 'Where are you living today?' They may or may not be in the same living situation they were in at the earliest relevant date. We need to decide which aspect we are really targeting here. It could be the same in cases of acute disease, but generally, we refer to various dates, and it typically defaults to the earliest date associated with the CCCD. There might be a more straightforward way to phrase this so that people understand what is intended to be captured.
  - Laura Erhart noted that all the elements for the CCCD would not be needed. The design behind it is that it will be defined for everyone, as everyone will have at least one of those elements. However, if we don't know the onset date, we simply don't have that information. The concept is intended to capture the earliest date related to that person's illness. For many conditions, this would likely start with the onset, but the CCCD is not meant to capture the beginning of the incubation period or the exposure period.
- Laura Erhart asked participants to visualize this for the poll:
  - A will represent the current living situation of the person at the time of the CCCD. We may switch out that wording, but that’s the concept.
  - B will refer to the current living situation of the person.
  - C will denote the current living situation of the person at the time of the report.
  - D will be for 'Other
- Steve Eichner responded that this wording is still a problem because 'current' implies a specific point in time. For example, if the report was made on January 1st, and we are now in July, what do we mean by 'current'?
  - Laura Erhart replied that the word “current” could be removed.
  - Steve Eichner agreed with that approach and noted that when they investigated the data collection for CCCD, they considered who was collecting the data and under what circumstances. Ultimately, it comes down to what the patient reports at that time, which is fundamental.
  - Laura Erhart commented that we will not reach a final decision on this today, but we can take the results from this poll and work on refinements.



Residence Type Value Set (Future Standard)

- Again, here is the value set we looked at before and had some discussion on them before polling. This value set is tied to the work being done with the Gravity Project and CDC's standardized tool for collecting this information in a case interview. It is a goal that these concepts will be captured in EHRs in the future.
- It includes a lot of detailed information about homelessness. For instance, 'private residence' and 'long-term arrangement' are listed at the top, followed by various other options.

Condensed Value Set

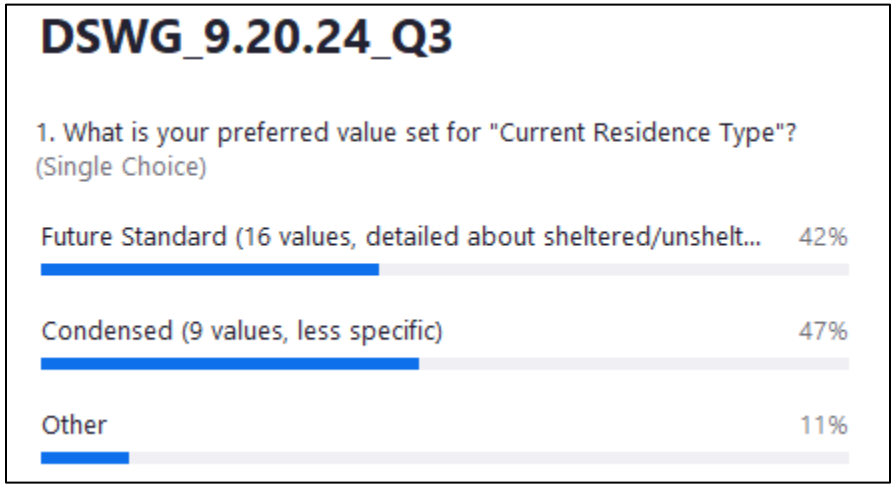
- This second value set is more condensed and captures fewer details about homelessness, so if the information is just coming from a medical record rather than a patient interview, you may have a better chance of having a value that represents what you know. It also provides a little more granularity on other congregate types of living situations that may be important to public health.

Condensed Value Set- Correctional Facility Details

- One piece of feedback received before this call is that the “type of correctional facility” data element in the TB MMG is useful because it allows for establishing a denominator for these populations. If this condensed value set is chosen in our first poll, we will then discuss the possibility of separating the option for “correctional facility” into the five options nested here.

Poll 3 (n= 38) – Residence Type Value Set

Given there was not a lot of time left for discussion, the poll was launched to get a pulse for participant's thoughts on this question.



Discussion

- Suzanne Gibbons-Burgener noted that they chose the 16 values because, in general, we prefer to be as granular as possible when collecting data. We can always condense it for analysis and creating indices, but once we condense the data, we can't go back.
- Misty Johnson commented that they chose the condensed version. If we're going to collect this as a core element across all diseases, it's better to keep it condensed. If knowing whether it's jail or prison makes a difference, we can add that as another question specific to that disease.
- Orlando Velazco noted that they preferred a condensed version. It was felt this core question is well-written, and that the condensed format better because it allows for follow-up. Since they are value sets, those responses can be expanded later. For instance, asking someone where they live can lead to follow-up questions like, 'Please describe what that means if they say they're in a shelter.' That's why Orlando opted for the condensed version; it's more universal and adaptable to different conditions.
- Sam Spoto noted that it's challenging for people to assess because the value sets are so lengthy, making it difficult to compare them side by side during a call like this. Additionally, they don't necessarily capture the same information. The longer value set could be condensed into those nine values, and in some respects, the condensed version offers more granularity depending on what aspect of the housing situation you're examining. I think it would be beneficial to create a Smartsheet poll and give everyone time to review the contents of both value sets. We could email it out to gather feedback, allowing people to fully consider the information and suggest improvements for either set. Regarding the condensed version and the distinction between homeless and congregate settings, such as a homeless shelter, should we consider whether to choose just one of these values or if we could implement a 'select all that apply' option to better capture the nuances of these situations?

        - Tyler Riedesel commented that they agreed with Sam. They voted for 'other' because they don't think either option is perfect or satisfactory, and I believe we can do better.
- Laura Erhart concluded by thanking everyone for the robust conversation.

Next Steps

- The SDOH DEST will receive a Smartsheet input form to continue to determine the best value set for “Residence Type.”
- The second Social Determinants of Health call will be held on October 18.
- The next Laboratory Reporting (Call #3) will be held on October 11.
- 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.