Data Standardization Work Group (DSWG)

April 26, 2024

Council of State and Territorial Epidemiologists

Please mute phone/ headset/ computer audio!

Agenda

- Welcome (Co-Chairs)
- DSWG Opportunities for Participation (Co-Chairs)
- Announcements and Key Updates (CSTE)
- Status Updates for closed DESTs (JMC)
- Review of Recent DEST Consensus (JMC):
  Exposures, Risk Factors, Underlying Conditions & Signs, Symptoms, and Clinical Manifestations
- Aligning Definitions with Person-Centered Language (JMC)
- DSWG Next Steps (JMC)
- Wrap-Up & Adjourn (Co-Chairs)

Welcome (Co-Chairs)

Core Data: In Scope for DSWG

- Identify core data elements for surveillance **for transmission to CDC**
  - For situational awareness or routine, cross-condition case surveillance
  - For case-based surveillance (not syndromic, not aggregate reporting)
  - Within *selected* data classes (e.g., hospitalization, travel history) that are used for multiple conditions
- Identify common definitions, value sets/code sets, structure, and interpretation notes for those core data elements
- Recommend best practices for how selected data elements are collected, structured or populated within PHA surveillance systems, if/when that will affect the interpretation of the data transmitted to CDC

Core for CDC Transmission ≠ Required

Even if a data element or class is identified as appropriate for transmission to CDC, this does *not* mean we need to fill in or send an answer for every core element for every case!

- It is okay to leave the data field as blank or Unknown if the information is not available, not a priority, or not useful for particular conditions.

Considerations

- What information is needed at the national level?
- How can we best take advantage of electronic records and automation?
- How can we best take advantage of and feed into other standardization efforts (e.g., HL7, USCDI, USCDI+)?
- How can we reduce the manual surveillance burden on public health and health care entities?
- Can we position ourselves for a more efficient process for adding emergent conditions to our systems? Exchanging information? Case notification onboarding?

Meeting ground rules

Discussion Guidelines

- "Raise" your hand
- Use the chat to provide feedback and ask questions
- When called, state your name and affiliation
- Stay on point, be brief, and within the context of the discussion in progress
- Co-Chairs/moderators' prerogative to move the discussion forward as needed
  - Post-meeting survey provided for additional input
- Be courteous and be open to new ideas

Active participation encouraged for all!

- Chats, voting, comments, etc.
- Includes STLT jurisdictions + CDC Programs + CSTE, etc.

DSWG Opportunities for Participation (Co-Chair)

CSTE

Data Element-Specific Teams

Purpose

These teams will focus on pre-identified topics determined by the DSWG Leadership Group (Co-Chairs, CSTE, CDC) based on multiple variables, including:

- Previous input from DSWG members
- Potential for collaboration across other groups of interest
- Common themes heard across various membership groups, including CSTE Sub-committees

Membership is encouraged from state/local/territorial jurisdictions, tribal entities, CDC and other federal agencies.

- To sign up fill in this interest form: https://app.smartsheet.com/b/form/3da89d48bda94fc7a33c4a8a505fe54a

Data Element-Specific Teams

Team Member Expectations

- Attend monthly Team meeting(s) regularly to discuss data element(s)
- Participate actively during and between meetings
- Provide feedback on how your jurisdiction or program (if applicable) collects and uses information related to the data element prior to the first Team meeting
    - All DSWG members are invited to participate in this step, but expectation is higher for Team members
- Respond to surveys/polls conducted between Team meetings
- Review and provide input and comments between meetings on any draft documentation using shared Word documents

There is **NO** expectation for Team members to write documents, lead meetings, or prepare slides.

Team Topics

- Upcoming Data Element-Specific Teams include:
  - Disability: **May 10th 1-2PM ET**
  - Identifiers: **May 17th 1-2PM ET**
    - Outbreak identifiers, Linkage of maternal/infant cases, Identifiers to link labs and cases
  - Upcoming: Race/Ethnicity/Tribal Affiliation, Social Determinants of Health (SDOH), Laboratory, Medication/Treatment, Reporting Source
- Individuals can opt-in to more than one team. Please keep in mind active participation is expected when participants sign up.
- Two Data Element-Specific Teams run concurrently. When discussions for one data elements are concluded, a new data element convenes in the existing meeting time.
- After guidance documents are developed for the data element topic, a team is sunset.

DSWG

Purpose

The larger DSWG Membership will convene bi-monthly, hear updates from Data Element-Specific Teams, and continue to provide feedback on written recommendations being developed by the DSWG.

DSWG meetings may also serve as a forum for discussion of broader topics or questions.

Member Responsibilities

- Attend meetings, as able, to hear updates and provide input to written recommendations
- Consider joining Data Element-Specific Teams
- Provide input on materials shared with the larger DSWG group in between meetings via polls or in shared documents
- Invite colleagues or other participants to relevant topic-specific sessions as able

Announcements and Key Updates (CSTE)

Announcements

CSTE 2024 Annual Conference

- Pittsburgh, Pennsylvania, June 9-13
- Early Bird Registration is now open through May 2nd: https://www.csteconference.org/index.php/registration/

- DSWG will have a roundtable/discussion session at the conference!
  - Surveillance & Informatics I - Data Standardization Workgroup (DSWG): Navigating Forward – Updates and Future Directions
  - Monday, June 10, 2024: 5:30 PM - 6:15 PM

Key Updates

- CSTE Best Practice Guideline – Almost Published (*like, next week*)!
  - *Data Standardization: Workgroup Proposal for Standardized Case Counting Date, Proposed Calculated Case Counting Date for NNDSS Case Counting*
- CSTE Brief - In Development (anticipated Spring 2024)
  - *Data Harmonization*

Status Updates for Closed DESTs (JMC)

Status update for closed DESTs

- DSWG review complete and draft is with CSTE for next steps.
  - Hospitalization
  - Travel
  - Pregnancy
  - Industry and Occupation
- DSWG input received, and final draft is being completed with Co-Chairs and JMC. The finalized draft will go to CSTE for next steps.
  - Immunizations
  - Death
  - Sexual Orientation and Gender Identity
- DEST members will soon receive a draft of the recommendations for feedback. (*Recent consensus will be reviewed on the following slides.)*
  - Exposures, Risk Factors, and Underlying Conditions
  - Signs, Symptoms, and Clinical Manifestations

DEST Summary (as of 4/26/2024)

Data Class	WG Discussion	DEST	Draft Complete	DEST Feedback	Feedback Addressed	DSWG Feedback	Document Finalized*
Hospitalization							
Travel							
Pregnancy							
Industry and Occupation							
Sexual Orientation and Gender Identity							
Immunization History							
Death							
Exposures/Risk Factors/ Underlying Conditions							
Signs/Symptoms/ Clinical Manifestations							
Disability		Call #1: 5/10/24					
Identifiers		Call #1: 5/17/24					

* Document is with CSTE for final approval.

Legend
- Completed
- In Progress
- Not applicable

Review of Recent DEST Consensus (JMC)

CSTE

Data Class Review

- Each data class → Many data elements
  - Exposures:
    Within an exposure period, did the person get a tattoo?
    Within an exposure period, did the person consume eggs?
  - Risk Factors:
    Is the person a healthcare worker?
    Has the person ever used tobacco?
  - Underlying Conditions:
    Does the person have a diagnosis of diabetes?
    Is the person positive for HIV?
  - Signs, Symptoms, Clinical Manifestations:
    Did the person experience fever?
    Did the person develop encephalitis?

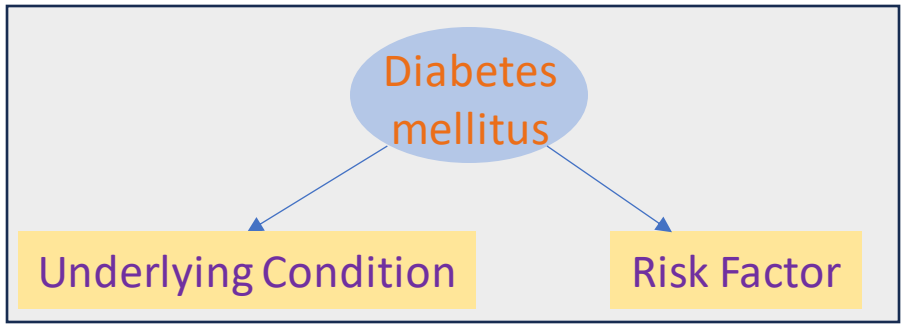
How will data elements be sorted into the data classes?

Data elements will be assigned to the data classes at the discretion of CDC programs.

→ Similar Data elements might be assigned to different data classes in different MMG's

Example:

Diabetes mellitus may remain as an Underlying Condition in the Respiratory and Invasive Bacterial Diseases (RIBD) MMG and may still be referred to as a Risk Factor in the COVID-19 mmg.



Data Class Structure Review

While at the fair did the person enter any of the following locations?

- [x] pig barn
- [ ] cattle barn
- [x] horse barn
- [ ] food tent
- [ ] restrooms

Repeating Element

While at the fair did the person enter any of the following locations?

pig barn

(•) Yes ( ) No ( ) Unknown

cattle barn

( ) Yes (•) No ( ) Unknown

horse barn

(•) Yes ( ) No ( ) Unknown

Repeating Group

While at the fair did the person enter the pig barn?

(•) Yes
( ) No
( ) Unknown

While at the fair did the person enter the cattle barn?

( ) Yes
(•) No
( ) Unknown

While at the fair did the person enter the horse barn?

(•) Yes
( ) No
( ) Unknown

Individual Questions

Data Class Structure Review

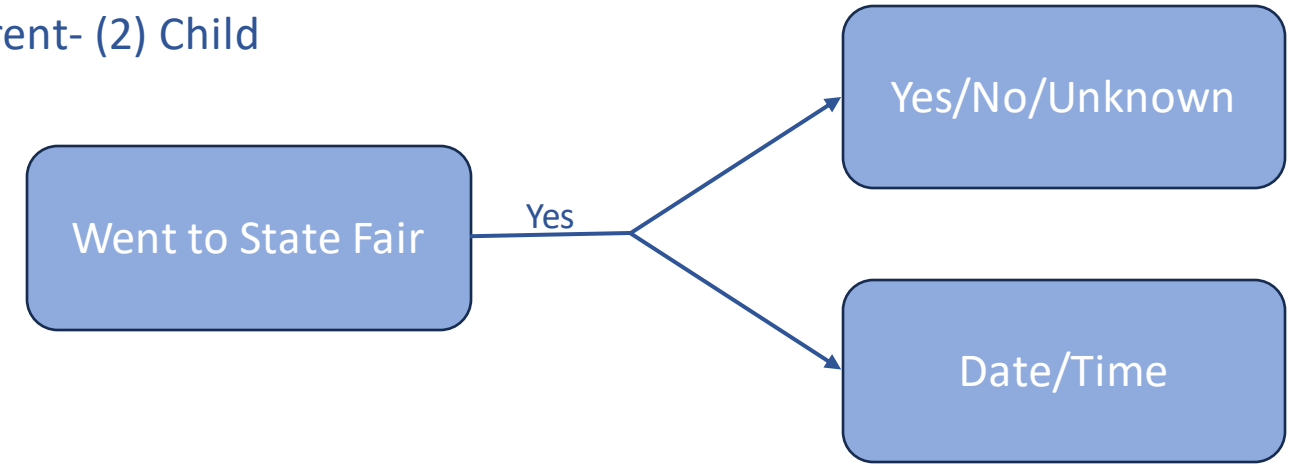
- Individual Questions
  - Require a full OMB review to clear each new question
  - Often straightforward to map and work into the system front-end to back
- Repeating Element
  - The lack of no/unknown can be less than ideal
  - Allows updating of value sets without OMB clearance process
  - Can be easily set up in surveillance systems and mapped to an HL7 message
- Repeating Group
  - Allows updating of value sets without OMB clearance process
  - Would allow for potential additional data points, for example a date to pair with each exposure or symptom
  - Can be quite an effort to map successfully
    - For some systems, subsequent mappings are easier once the process is figured out. For others, each mapping is resource intensive even when the initial process is understood.

Data Class Structure Review

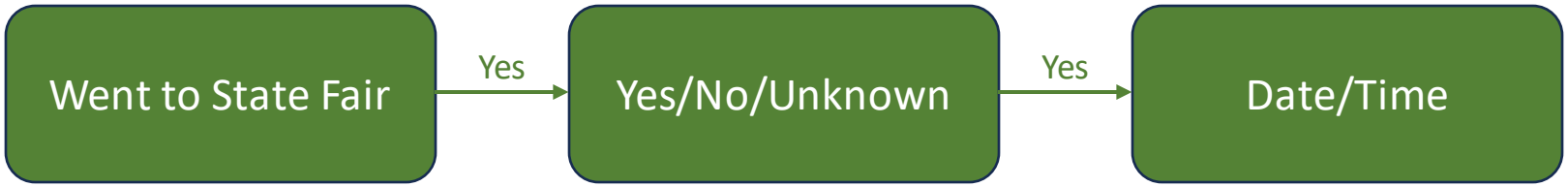
- Structure should be what makes sense for the data class.
- CDC: strong preference for repeating group
- STLTs: preferred structure was split between the three, likely reflecting differences in surveillance systems used across the US
- Adding additional elements into a repeating group can be complex to implement in some jurisdictional systems.
  - Parent/child/child is simpler than parent/child/grandchild*
- Strong message that consistency in structure across conditions for a data class is important.

Data Class Structure Review

Parent- (2) Child



Parent-Child-Grandchild



Exposures, Risk Factors, and Underlying Conditions

CSTE

ERU Proposed Definitions

Data Class	Suggested Definition	Examples
Exposures	Within exposure window, either the person's contact with an entity (person, animal, or substance) or presence at a location where exposure to an agent could have occurred. (*e.g.* where did pathogen come from).	Within an exposure period: • Visited a water park • Got a tattoo • Worked in a daycare
Risk Factors	Characteristics, behaviors, or social factors that may increase the likelihood of developing a disease or condition.	• Substance Use ever • Unstable Housing • Smoking Status
Underlying Health Conditions	Pre-existing health conditions that may worsen the course or severity of a new disease or health condition.	• COPD • Diabetes • Prior History of Disease

ERU Data Class Indicators

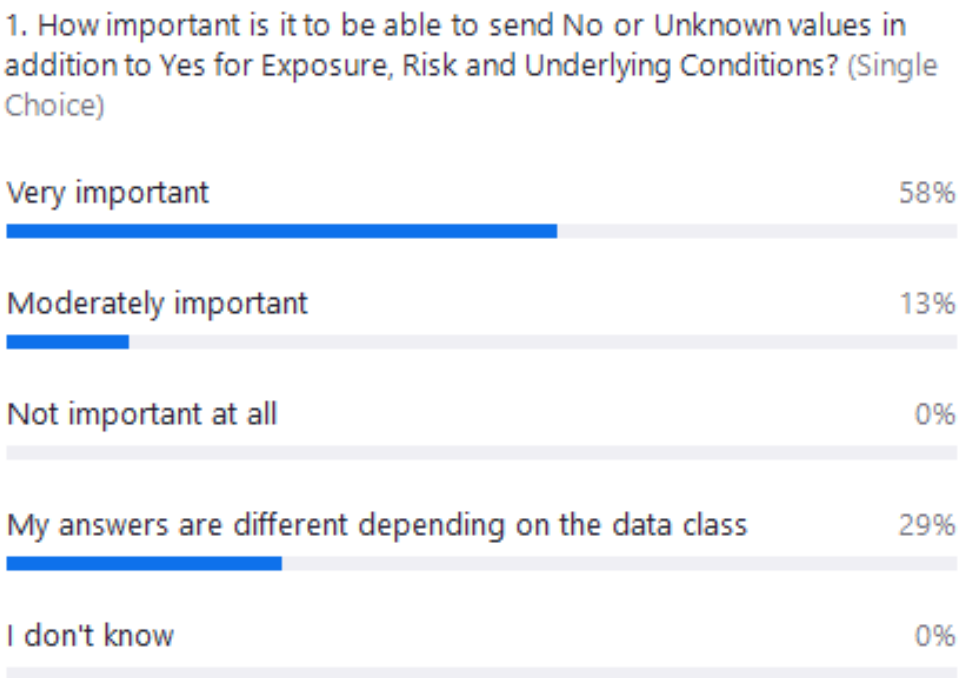
Data Element (DE) Name	Data Element Description	May Repe	Value Set Name (VADS Hyperlink	HL7 Cardinality	HL7 Implementation Note	Repeating Group Eleme
Did Underlying Condition(s) Exist	Did the subject have any underlying causes or prior illnesses?	N	Yes No Unknown (YNU)	[0..1]		NO

- Answering "No" may remove the need to send the repeating group or explain why a repeating element is absent.
- May already exist in some systems.
- This would be a distinct summative data element for each data class.

Data Concept	Suggested Definition	Example Question
Exposure Indicator	Indicates if the person experienced any exposure included in the condition exposure value set.	Did the person experience any of the exposures in the list provided?

Proposed Structure Recommendations

- Structure should be what makes sense for the data class.
- Consistency in structure across conditions for a data class is very important.
- Repeating element should likely not be used for any of these data classes.



Signs, Symptoms, and Clinical Manifestations

CSTE

Potential data elements related to Signs, Symptoms, and Clinical Manifestations

- Signs and Symptoms Repeating Group
- Clinical Manifestations Repeating Group
- Clinical Findings Repeating Group
- Problem Associated Signs and Symptoms
- Clinical Signs of Congenital Syphilis
- Symptoms Resolved After Treatment
- Late Clinical Manifestations
- Complications
- Infection Type

Signs and Symptoms - Clinical Manifestations

- Signs and Symptoms
  - Carbon Monoxide Poisoning, Rubella, Measles, RIBD, Congenital Syphilis, Foodborne and Diarrheal Disease, Arboviral, Listeriosis, Mumps, Pertussis, COVID-19

Data Element (DE) Name	DE Identifier Sent in HL7 Message	Data Element Description	Value Set Name (VADS Hyperlink)
START: Signs and Symptoms Repeating Group			
Signs and Symptoms	56831-1	Signs and symptoms associated with the illness being reported	Signs and Symptoms (COVID-19)
Signs and Symptoms Indicator	INV919	Indicator for associated sign and symptom	Yes No Unknown (YNU)
END: Signs and Symptoms Repeating Group			

- Clinical manifestations
  - Babesiosis, Brucellosis, Leptospirosis, Lyme Disease, TBRD

Proposed Data Elements

Most MMGs utilize either a "signs and symptoms" repeating group or "clinical manifestations" repeating group.

This data class intends to encompass/standardize those.

Inclusive of signs, symptoms, and other clinical manifestations like an "abnormal chest x-ray" or "low white blood cell count".

Data Element	Definition	Code System or Value Set
Clinical Manifestations	Clinical manifestations, including signs and symptoms, associated with the condition being reported	SNOMED preferred, local if necessary

Proposed Data Elements

- May already exist in some systems.
- This would be a distinct summative data element that conveys if any of the values within "clinical manifestations" are yes, if all of them are "no", or if all are "unknown".
- Answering “no” or "unknown" may remove the need to send individual responses.

Data Element	Definition	Code System or Value Set
Summative Clinical Manifestations Indicator	**Indicator that the person experienced any clinical manifestations associated with the condition being reported**	**Y/N/U**
Clinical Manifestations	Clinical manifestations, including signs and symptoms, associated with the condition being reported	SNOMED preferred, local if necessary

Proposed Data Elements

- Conveying if the person has any clinical manifestations is important epidemiologically, and helpful for mapping of data for some jurisdictions.
- Part of the Minimum Dataset for Emergency Use is a "symptomatic" indicator- a patient may have a clinical manifestation but perceive themselves as "asymptomatic".
- Perception of symptoms may be more important for public communication.
- Both indicator values were voted as core (weak support) by the DEST.

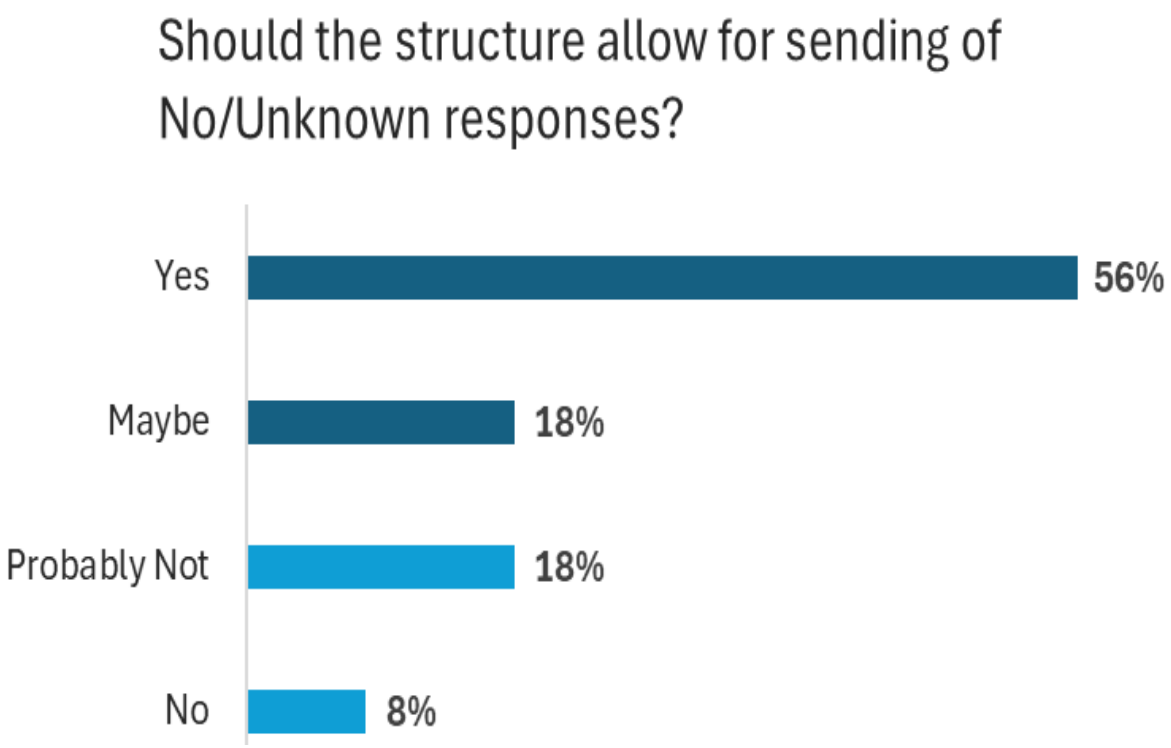
Data Element	Definition	Code System or Value Set
Symptomatic Indicator	Indicator that the person experienced any symptoms associated with the condition being reported	Y/N/U

Complications - Infection Type - Clinical Findings

- Complications
  - FDD: Hemolytic Uremic Syndrome, Thrombotic Thrombocytopenia
  - Malaria: Cerebral Malaria, Renal Failure, etc.
  - Babesiosis: Acute Respiratory Distress, Congestive Heart Failure, etc.
  - Mumps: Deafness, Mumps Encephalitis, Mumps Orchitis, etc.
  - Pertussis: Encephalopathy, Seizures, Other
  - Varicella: Hemorrhagic Condition, Varicella Encephalitis, etc.
- Infection Type
  - RIBD: Describes type of infection caused by the organism (Cellulitis, Encephalitis, Septic Shock, etc.)
- Clinical Findings
  - COVID-19: Acute Respiratory Distress Syndrome, Electrocardiogram Abnormal, Pneumonia, etc.

Proposed Structure Recommendations

- Structure should be what makes sense for the data class.
- Consistency in structure across conditions for a data class is very important.
- Repeating element should not be used for this data class.



Aligning Definitions with Person-Centered Language (JMC)

CSTE

Language in Previous Definitions

- How is a person referred to across our definitions so far?

Hospitalization	Subject, Patient
Travel	Patient
Pregnancy	Case
I/O	Subject
Sex	Person
Death	Patient
Immunizations	Patient
ERU	Person
SSCM	Person

Updating Language

- Let's take this opportunity to align the wording used across definitions, when appropriate.
    - Patient: As public health surveillance, we are not in the role of providing medical care or treatment
    - Subject, Case: Consciously or subconsciously focuses on the condition being reported rather than person with the condition
    - Person: The data being transmitted is always about a person

Updating Language

Hospitalization	**Admission Date**	The date on which the ~~subject~~ **person** was admitted
Hospitalization	**Discharge Date**	The date on which the ~~subject~~ **person** was discharged
Hospitalization	**Admitted to ICU**	Indicator that the ~~patient~~ **person** was admitted to an intensive care unit (ICU) or other specialty care unit for the condition

Updating Language

Travel	**Travel**	Indicator that the ~~patient~~ **person** spent time outside of their jurisdiction of residence in a time period relevant to the condition
Travel	**International Travel**	Indicator that ~~patient~~ **person** spent time outside of the USA in a time period relevant to the condition
Travel	**Domestic Travel**	Indicator that the ~~patient~~ **person** spent time domestically outside of their jurisdiction of residence, but within the USA, in a time period relevant to the condition
Travel	**International Destination(s) of Recent Travel**	International destination or countries the ~~patient~~ **person** spent time within

Updating Language

Pregnancy	**Pregnancy Status**	Pregnancy status of the ~~case~~ **person** at the time of the event (CCCD*)
Death	**~~Patient~~ Person Deceased**	Indicator that the ~~patient~~ **person** is deceased.
Death	**Deceased Date**	~~Patient~~ **Person's** date of death
Immunizations	**Ever Vaccinated Against this Disease**	Indicator that the ~~patient~~ **person** has ever received a vaccine against this disease.

Updating Language

I/O	Current Industry	Narrative text of a ~~subject's~~ **person's** current industry
I/O	Current Industry Standardized	The CDC NIOSH standard industry code and standard text based upon the narrative text of a ~~subject's~~ **person's** current industry.
I/O	Current Occupation	Narrative text of a ~~subject's~~ **person's** current occupation
I/O	Current Occupation Standardized	The CDC NIOSH standard occupation code and standard text based upon the narrative text of a ~~subject's~~ **person's** current occupation.

Person-Centered Language Approach

We are planning to update all mentions of subject/case/patient to the word "person" in our definitions.

Table 6. Core Surveillance Data Element for R e CDC Case Notification – Ever Vaccinated Against this Disease

Ever Vaccinate gainst this Disease	
Definition	Indicator that the patient has ever received a vaccine against this disease.
Standard Code for Transmission to CDC	PHIN Questions: VAC126 – Did the Subject Ever Receive a Vaccine Against This Disease
Value Set or Coding	Yes/No/Unknown

Does anyone have concerns about this plan?

DSWG Next Steps (JMC)

CSTE

Next Steps

- In early May, DEST members should look for an email with the draft Exposures, Risk Factors, and Underlying Conditions & Signs, Symptoms, and Clinical Manifestations recommendation documents and feedback forms.

- If interested, sign up for the next data classes of Identifiers and Disability
  - Identifiers will include discussions of outbreak IDs, methods to link maternal and infant cases, and methods to link laboratory results with cases

Other Next Steps

Upcoming meetings

- Disability: **May 10th 1-2PM ET**
- Identifiers: **May 17th 1-2PM ET**
  - Outbreak identifiers, Linkage of maternal/infant cases, Identifiers to link labs and cases
- Future data classes:
  - Race/Ethnicity/Tribal Affiliation
  - Social Determinants of Health (SDOH)
  - Laboratory
  - Medication/Treatment
  - Reporting Source

The next broad DSWG call will be **June 28, 1-2 ET**.

Wrap-Up & Adjourn (Co-Chairs)

Anything Else?