Core Case Notification Data Element Recommendations: Social Determinants of Health

This document was developed by the Data Standardization Work Group (DSWG) of the Council of State and Territorial Epidemiologists (CSTE) Surveillance Practice and Implementation Subcommittee (SPIS) as part of a larger data harmonization effort. It will be published and maintained in a repository on the CSTE website.

The DSWG identified and standardized the data elements to utilize in the transmission of core surveillance case data to the Centers for Disease Control and Prevention (CDC). Core surveillance data classes and data elements are those that apply across many conditions and, in general, are not specific to any single condition or condition family. In each recommendation document, the DSWG provides common definitions, value sets/code sets, structure, and interpretation notes for each identified data element that is designated as core. The DSWG does not provide questions which would be used to populate the recommended data elements. More information about the DSWG and data harmonization can be found here.

The structure of this document is as follows:

Background

In public health frameworks, a distinction has been made between Social Determinants of Health (SDOH) at the community level and health-related social needs at an individual level.[1] Although the naming convention for this data class is the well-recognized term "SDOH", the data elements defined in this document represent health-related social needs since the data collected at the case level are person-level data. These are “social and economic needs that individuals experience that affect their ability to maintain their health and well-being", as opposed to SDOH that reflect community-level conditions in the environments where people live that affect their health. [2] The data elements discussed below help public health entities identify which populations may be at higher risk of adverse health outcomes due to their social conditions.

Standardization efforts by the DSWG focused on five topic areas: residence type, primary/preferred language, education, income, and insurance. Although other topics relate to SDOH, these areas were chosen based on an initial poll of group members and prioritized to **capture snapshot demographic data elements for a person at a time relevant to the condition being reported. These data elements are meant to capture data for analysis of social factors affecting health** rather than capturing a person’s exposures or risk factors. Additional information related to SDOH data elements may also be captured in the Exposures or Risk Factors data classes *[link to ERU document in final version]* to convey connected but distinct concepts such as “during the exposure window did the person live in a dormitory?” or “has the person ever been incarcerated?”

[1] Addressing Health-Related Social Needs in Communities Across the Nation | U.S. Department of Health and Human Services. (Accessed January 9, 2025). https://aspe.hhs.gov/sites/default/files/documents/3e2f6140d0087435cc6832bf8cf32618/hhs-call-to-action-health-related-social-needs.pdf

[2] Social Drivers of Health and Health-Related Social Needs | Centers for Medicare and Medicaid Services. (Accessed August 21, 2024). https://www.cms.gov/priorities/innovation/key-concepts/social-drivers-health-and-health-related-social-needs

During the consensus building process, the DSWG considered previous work conducted by the Gravity Project, a "national public collaborative that develops consensus-based data standards to improve how we use and share information on SDOH."[3] The group has identified value sets and corresponding terminology (e.g., SNOMED and LOINC codes) that enable meaningful exchange of standardized information between data systems (e.g., between public health data systems or on a health information exchange). Definitions identified by the Gravity Project membership for selected data elements applicable to this DSWG effort are shown in Table 1. Work related to language barriers and incarceration status is currently in progress.

Table 1. A subset of Gravity Project data elements and associated definitions.

Data Element	Gravity Project Definition
Sheltered Homelessness	Currently living in a shelter, motel, temporary or transitional living situation, scattered-site housing, or not having a consistent place to sleep at night
Unsheltered Homelessness	Residing in a place not meant for human habitation, such as cars, parks, sidewalks, abandoned buildings, on the street
Less than High School Education	Failing to meet academic criteria for high school diploma or equivalent
Financial Insecurity	A state of being wherein a person has difficulty fully meeting current and/or ongoing financial obligations and/or does not feel secure in their financial future
Health Insurance Coverage Status	Documentation of presence and type of insurance

[page left intentionally blank]

[3] Our Mission | Gravity Project. (Accessed January 9, 2025). https://thegravityproject.net/overview/

Core Surveillance SDOH Data Elements for Routine Case Notifications

The following data elements were considered by the DSWG membership to be core surveillance data under the umbrella of the SDOH data class. Data elements marked with an asterisk (*) did not reach the 80% consensus mark within the group but were supported by more than half of the participants. These are included here as core data elements with weak support from the DSWG. For each SDOH-related data element, there is no pre-defined default value for what should be transmitted when the data element has not been populated. Refer to the following sections for a background discussion of each topic, and Tables 4, 5, 7, 10 and 11 below for definitions and additional information.

- Housing Status
- Primary Living Situation
- Preferred Language
- Educational Attainment
- Healthcare Insurance or Access Type*

During compilation of background research, CDC did not identify any National Notifiable Diseases Surveillance System (NNDSS) program areas that have a use case for the collection of income. Therefore, income was not discussed as a potential core data element.

During state epidemiologist review and voting for this data class, the data elements Primary Living Situation, Preferred Language, Educational Attainment, and Healthcare Insurance or Access Type received a significant (although minority) portion of votes that do not support them as core data elements (30-40%) and will not support implementation (37%) of this data class. Many comments were received that elaborate on how these data elements are nice to have but not essential for public health action, may be more important at the local level than the federal level, add burden in data collection for investigators, and may cause further mistrust and lack of participation with public health. In addition to these reasons, many of these data elements are unable or unlikely to be transmitted via ELR, eCR, or present in a medical record, therefore many cases will not have these data elements populated. While the DSWG does **not** define a core data element as “required” for transmission to CDC for any data class, it is imperative to address the above concerns by strongly recommending these data elements be implemented as low priority for transmission and not required to be populated.

[page left intentionally blank]

Residence Type

Current CDC Message Mapping Guides (MMGs) for case notifications sent to the National Notifiable Diseases Surveillance System (NNDSS) contain various data elements related to residence and homelessness (Table 2). Standardization of a Residence Type data element as it relates to SDOH is not intended to replace all use cases listed in Table 2, as many assess residence as it relates to exposure to a condition or as a historical risk factor.

Table 2. Data elements related to residence and homelessness from CDC Message Mapping Guides (MMGs).

MMG	Data Element Name	Identifier	Data Element Description	Value Set Name
Foodborne and Diarrheal Disease	Homeless	32911000	No fixed residence for any given period of time	Yes No Unknown (YNU)
COVID-19	Residence	75617-1	Which would best describe where the patient was staying at the time of illness onset?	Residence Type (COVID-19)
Respiratory and Invasive Bacterial Disease (RIBD)	Residence	75617-1	Where was the patient a resident at time of initial culture?	Residence Location (RIBD)
Respiratory and Invasive Bacterial Disease (RIBD)	Homeless	32911000	Was the patient homeless at time of symptom onset?	Yes No Unknown (YNU)
Hepatitis	Type of Incarceration Facility	INV638	Type of facility where the patient was incarcerated for the longer than 24 hours before symptom onset?	Incarceration Type (Hepatitis)
Hepatitis	Incarcerated more than 24 hours	INV637	Prior to the onset of symptoms, was the patient incarcerated for longer than 24 hours? For Acute Hep B, the time period prior to onset of symptoms is 6 weeks - 6 months. For Acute Hep C, the time period prior to onset of symptoms is 2 weeks - 6 months.	Yes No Unknown (YNU)
Hepatitis	Incarcerated more than 6 months	INV639	Was the patient ever incarcerated for longer than six months during his or her lifetime?	Yes No Unknown (YNU)
Hepatitis	Ever Incarcerated	INV649	Was the patient ever incarcerated?	Yes No Unknown (YNU)
STD	Been incarcerated within the past 12 months	STD118	Been incarcerated within the past 12 months	YNRD (CDC)
Tuberculosis and Latent TB	Type of Correctional Facility	INV1119	If patient was a Resident of Correctional Facility at Diagnostic	PHVS_CorrectionalFacilityType_NND

			Evaluation, indicate the type of correctional facility.	

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) and residence in a congregate setting. A person would be asked to provide information about where they are living in their own words, and the investigator would then 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 is working with the Gravity Project to develop standards and value sets related to this question set.

Residence in a correctional facility can have impact on a person's health through factors like overcrowding, limited access to healthcare, or systemic and social obstacles with family, employment, housing, and health upon release.[4] The Bureau of Justice statistics collects and provides data on adult correctional populations in the US, including population estimates for relevant correctional facility types. The Office of Juvenile Justice and Delinquency Prevention provides one-day census estimates for facilities identified as juvenile residential placement facilities. Applicable definitions utilized by these agencies are captured in Table 3.

Table 3. A subset of relevant definitions used by the Bureau of Justice Statistics and Office of Juvenile Justice and Delinquency Prevention

Term	Definition
Adult[5]	A person subject to the jurisdiction of an adult criminal court or correctional agency. Adults are those age 18 or older in most jurisdictions. Persons age 17 or younger who were prosecuted in criminal court as if they were adults are counted as adults, but persons age 17 or younger who were under the jurisdiction of a juvenile court or agency are excluded. Local jails, however, may hold persons age 17 or younger before or after they are adjudicated.
Local jail population[5]	Estimated number of inmates held in confinement facilities operated under the authority of a sheriff, police chief, or city or county administrator. Facilities are intended for adults but may hold juveniles before or after they are adjudicated. Facilities include jails, detention centers, city or county correctional centers, special jail facilities (such as medical or treatment centers and prerelease centers) and temporary holding or lockup facilities that are part of the jail's combined function. Inmates sentenced to jail facilities usually have a sentence of 1 year or less.
Prison Population[5]	Estimated number of prisoners incarcerated in a long-term confinement facility run by a state or the federal government and typically holding felons and other persons with sentences of more than 1 year, although sentence length may vary by jurisdiction.
Census of Juveniles in Residential	CJRP asks all juvenile residential facilities in the United States to describe each person younger than 21 assigned a bed in the facility on the census date as a result of a status or delinquency offense. CJRP does not capture data on youth in adult prisons or jails or those placed in facilities used exclusively for mental health or substance use treatment or for dependent children.

[4] Incarceration | Healthy People 2030. (Accessed January 17, 2025.) https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/incarceration

[5] Correctional Populations in the United States, 2022 – Statistical Tables | U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics. (Accessed January 17, 2025.) https://bjs.ojp.gov/document/cpus22st.pdf

Placement (CJRP)[6]	

While one data element of “Residence Type” was initially discussed, the DSWG consensus was that residence type is better transmitted as two distinct data elements of Housing Status, which captures if someone is experiencing homelessness or unhoused, and then separately Primary Living Situation, which is a description of where someone lives (e.g., private residence, correctional facility, or long-term care facility). This is not intended to capture if someone has ever experienced homelessness or been incarcerated (which could be captured in the Risk Factors data class) or if someone experienced homelessness or incarceration at any point during the incubation period for the condition (which could be captured in the Exposures data class). The DSWG aimed to align the value set for Primary Living Situation with the work being done by CDC, Fulton County, and the Gravity Project.

At the federal level multiple definitions exist for homelessness, complicating data collection, analysis, and comparison. The US Department of Housing and Urban Development’s (HUD) Point-in-Time (PIT) Count is a commonly used dataset for estimating the burden of sheltered and unsheltered homelessness in the US. The methodology directs the definitions utilized in the PIT Count as[7]:

- ***Sheltered:*** *An individual or family living in a supervised publicly or privately operated shelter designated to provide temporary living arrangement (including congregate shelters, transitional housing, and hotels and motels paid for by charitable organizations or by federal, state, or local government programs for low-income individuals).*
- ***Unsheltered:*** *An individual or family with a primary nighttime residence that is a public or private place not designed for or ordinarily used as a regular sleeping accommodation for human beings, including a car, park, abandoned building, bus or train station, airport, or camping ground.*

These definitions are similar to those used by the Gravity Project in Table 1, although the Gravity Project includes “not having a consistent place to sleep at night” in the definition of sheltered homelessness that is not represented in the HUD definition above. The DSWG discussion recognized the importance of using a standardized definition for homelessness for data analysis while also recognizing that during case investigation it may be identified that a person is experiencing homelessness without documentation or knowledge of their specific living situation. To achieve this goal, the concept of residence captured as two data elements allows for a less specific documentation of a person’s "Housing Status” (and therefore potential health-related social needs) (Table 4), as well as a more specific “Primary Living Situation” (Table 5) aligning with the categories defined by HUD PIT Count methodology for purposes of data analysis.

For the Primary Living Situation value set, the DSWG discussed including a single option for “correctional facility”, versus the five options present in value set in Table 5. The group found it important to capture the type of correctional facility where the individual is residing as different facilities represent distinct populations of public health importance. The correctional facility value used should indicate the type of facility where a person resides, regardless of the type of agency that may have legal custody over the person (e.g., for someone in state custody who is temporarily residing in a local jail, the “local detention facility” value should be used).

[page intentionally left blank]

[6] Census of Juveniles in Residential Placement | U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention. (Accessed January 17, 2025.) https://ojjdp.ojp.gov/research-and-statistics/research-projects/Census-of-Juveniles-in-Residential-Placement/overview

[7] Point-in-Time Count Methodology Guide. | U.S. Department of Housing and Urban Development. 2014. (Accessed February 11, 2025). https://files.hudexchange.info/resources/documents/PIT-Count-Methodology-Guide.pdf

Table 4. Core Surveillance Data Element for Routine CDC Case Notification – Housing Status

Housing Status	
Definition	Housing status of a person at the first epidemiologically relevant date associated with the case (i.e., CCCD*)
Standard Code for Transmission to CDC	LOINC: 71802-3 Housing status
Value set or coding system	32911000 | Homeless (finding) 414418009 | Housed (finding) 261665006 | Unknown (qualifier value)
Abstracted/ Derived/ Created	Abstracted
Recommendations for Input Data Source(s)	• Case investigation • eCR • ELR
Notes	This data element is a snapshot of housing status at the earliest epidemiologically relevant date associated with the case (i.e., CCCD) to assess it as a health-related social need consistently across conditions. This is not intended to capture if someone has ever experienced homelessness (which could be captured in the Risk Factors data class) or if someone experienced homelessness at any point during the incubation period for the condition (which could be captured in the Exposures data class). The use of "32911000 | Homeless (finding)" may represent those meeting either the Gravity Project or HUD PIT Count definitions, as well as documentation of unspecific homelessness where the specific living situation is unknown. The data element "Primary Living Situation" allows for data analysis aligning with the categories specifically defined by denominator data, such as the HUD PIT Count methodology.
Data Source Details	
eCR CDA	HL7 CDA R2 Implementation Guide: Public Health Case Report, Release 2 STU Release 1.1 - US Realm • A standardized method of transmitting housing status is not outlined, but relevant information may be found as a Social History Observation within the Social History Section. • ClinicalDocument/component/Social History Section/Social History Observation/observation/value HL7 CDA R2 Implementation Guide: Public Health Case Report – eICR Release 2, STU Release 3.1 - US Realm • ClinicalDocument/component/Social History Section/Characteristics of Home Environment/observation/value • The implementation guide specifies that the value transmitted within Characteristics of Home Environment for someone who is unhoused should be the SNOMED CT code "32911000 |Homeless (finding)".
eCR FHIR	Resource Profile: US Public Health Characteristics of Home Environment • Observation.value[x] = Unless not suitable, codes SHALL be taken from Residence and Accommodation Type • Value set Includes "32911000 |Homeless (finding)".
ELR	No field in ELR directly maps to this data element, but relevant information may be found in the Patient Address field.**
USCDI v6	No direct comparison to this data element exists, but relevant information may be found in Current Address** or the following:

	SDOH Assessment • Definition: Screening questionnaire-based, structured evaluation for a Social Determinants of Health-related risk. Examples include but are not limited to food, housing, transportation security, and health literacy. • Applicable vocabulary: LOINC, SNOMED CT • Part of Health Status Assessments data class SDOH Problems/Health Concerns • Social Determinants of Health-related concerns, conditions, or diagnoses. Examples include but are not limited to homelessness and food insecurity. • Applicable vocabulary: SNOMED CT, ICD-10-CM • Part of the Problems data class
USCDI+ Public Health: Case Reporting Use Case	N/A

* The Calculated Case Counting Date (CCCD) is the earliest date available of six component dates and is referenced as the earliest epidemiologically relevant date associated with the case. Data Standardization: Workgroup Proposal for Standardized Case Counting Date

** Additional information on input data sources can be found in Appendix A.

[page intentionally left blank]

Table 5. Core Surveillance Data Element for Routine CDC Case Notification – Primary Living Situation

Primary Living Situation	
Definition	Primary living situation of a person at the first epidemiologically relevant date associated with the case (i.e., CCCD)
Standard Code for Transmission to CDC	LOINC: TBD
Value set or coding system	• Private residence in a long-term arrangement of more than 14 days • Hotel/motel or vacation rental in a long-term arrangement of more than 14 days • Private residence in a short-term arrangement of 14 days or less • Hotel/motel or vacation rental, short-term arrangement 14 days or less • Shelter or safe haven (congregate setting) • Temporary, non-congregate housing provided by charity or government program (e.g., transitional housing, pallet shelter, hotel/motel) • Structure or vehicle not meant for human habitation • Vehicle meant for human habitation (e.g., RV) • Open air • Migrant worker housing • Military barracks • School/university dormitory • Other congregate housing for workers/students • Federal adult correctional facility • State adult correctional facility • Local adult jail/detention facility • Juvenile correctional/detention facility • Other correctional/detention facility (e.g., border detention facility) • Mental/Behavioral/Substance use treatment facility • Long term care facility • Medical respite/Recuperative care facility • Group home or residential facility not provided by employer or school • Unknown • Other
Abstracted/ Derived/ Created	Abstracted
Recommendations for Input Data Source(s)	• Case investigation • eCR • ELR
Notes	The intent of this data element is to capture primary living situation (e.g., someone on vacation may have a primary living situation of a private residence even though they are currently in a short-term hotel stay). This data element is a snapshot of someone's primary living situation at the earliest epidemiologically relevant date associated with the case (i.e., CCCD) to capture this health-related social need consistently across conditions. This is not intended to capture if someone has ever lived in a certain situation (which could be captured in the Risk Factors data class) or all living situations at any point during the incubation period for the condition (which could be captured in the Exposures data class). Responses for a correctional facility should indicate the type of facility where a person resides, regardless of the type of agency that may have legal

	custody over the person (e.g. someone in state custody may temporarily reside in a local jail).
Data Source Details	
eCR CDA	HL7 CDA R2 Implementation Guide: Public Health Case Report, Release 2 STU Release 1.1 - US Realm • A standardized method of transmitting primary living situation is not outlined, but relevant information may be found as a Social History Observation within the Social History Section. • ClinicalDocument/component/Social History Section/Social History Observation/observation/value HL7 CDA R2 Implementation Guide: Public Health Case Report – eICR Release 2, STU Release 3.1 - US Realm • The value set Residence and Accommodation Type (contains 282 values) referenced by the template Characteristics of Home Environment within the Social History Section may contain relevant information to ascertain primary living situation. • ClinicalDocument/component/Social History Section/Characteristics of Home Environment/observation/value
eCR FHIR	Resource Profile: US Public Health Characteristics of Home Environment • Observation.value[x] = Unless not suitable, codes SHALL be taken from Residence and Accommodation Type • This value set may contain relevant information to ascertain primary living situation.
ELR	No field in ELR directly maps to this data element, but relevant information may be found in the Patient Address field.**
USCDI v6	No direct comparison to this data element exists, but relevant information may be found in Current Address** or the following: SDOH Assessment • Definition: Screening questionnaire-based, structured evaluation for a Social Determinants of Health-related risk. Examples include but are not limited to food, housing, transportation security, and health literacy. • Applicable vocabulary: LOINC, SNOMED CT • Part of Health Status Assessments data class SDOH Problems/Health Concerns • Social Determinants of Health-related concerns, conditions, or diagnoses. Examples include but are not limited to homelessness and food insecurity. • Applicable vocabulary: SNOMED CT, ICD-10-CM • Part of the Problems data class
USCDI+ Public Health: Case Reporting Use Case	No direct comparison to this data element exists, but “Congregate Living” is a data element within USCDI+. This concept is captured in some of the values within the Primary Living Situation value set. Congregate Living • Definition: Shared housing includes a broad range of settings, such as apartments, condominiums, student or faculty housing, national and state park staff housing, transitional housing, and domestic violence and abuse shelters.

	• Part of the Social Determinants of Health data class

* The Calculated Case Counting Date (CCCD) is the earliest date available of six component dates and is referenced as the earliest epidemiologically relevant date associated with the case. Data Standardization: Workgroup Proposal for Standardized Case Counting Date

**Additional information on input data sources can be found in Appendix A.

[page intentionally left blank]

Preferred Language

Collection of data on language is important for public health response and communication. It allows for effective communication of health materials in a language that individuals can understand, identifies language barriers that might affect access to healthcare services, and allows public health to offer services or materials that are culturally or linguistically appropriate. The Joint Commission that validates and accredits health care organizations and programs in the U.S. requires collection and documentation of the patient's preferred language (defined as the language the patient would like to use when discussing health care) to enable access for patients who speak a language other than English. Similarly, CDC's National Healthcare Safety Network (NHSN) currently allows optional documentation of the "preferred language(s) spoken by patients" and the need for medical interpretation services. NHSN has suggested all EHR vendors use the standard accepted ISO 639 codes for living languages and, at minimum, provide the NHSN Abridged Primary Language List of over 500 languages that includes more common languages spoken in the U.S. Other languages can be added to this list based on local geographic language needs, using the ISO 639 master list.

Two MMGs currently capture a data element of Primary Language - Carbon Monoxide (CO) and COVID-19 (Table 6).

Table 6. Data elements related to language from CDC Message Mapping Guides (MMGs).

MMG	Data Element Name	Identifier	Data Element Description	Value Set Name
Carbon Monoxide	Primary Language	DEM142	What is the patient's primary language?	Language
COVID-19	Primary Language	DEM142	What is the patient's primary language? Please indicate for both hospitalized and not hospitalized cases.	Language

The DSWG defines the core data element as "the language a person is most comfortable using for communication," intending to be widely encompassing. Similar to NHSN, the United States Core Data for Interoperability (USCDI) and electronic case reporting (eCR) define preferred language as the "language the patient would like to use when discussing healthcare." The language documented from this field in a medical record or eCR may be utilized if available; however, it has limitations as the perceived availability of translation services or societal stigma may influence a patient's answer to this question during a healthcare encounter.

[page intentionally left blank]

Table 7. Core Surveillance Data Element for Routine CDC Case Notification – Preferred Language

Preferred Language	
Definition	The language a person is most comfortable using for communication
Standard Code for Transmission to CDC	LOINC: 54899-0 Preferred language
Value set or coding system	Minimum Value Set: NHSN Abridged Primary Language List Additional codes from the ISO 639-3 code set can be utilized as needed.
Abstracted/ Derived/ Created	Abstracted
Recommendations for Input Data Source(s)	• Case investigation • eCR • ELR
Notes	eCR defines preferred language as the "language the patient would like to use when discussing healthcare". The information from this field in eCR may be utilized if available; however, it has limitations as the perceived availability of translation services or societal stigma may influence a patient's answer to this question during a healthcare encounter.
Data Source Details	
eCR CDA	HL7 CDA R2 Implementation Guide: Public Health Case Report, Release 2 STU Release 1.1 - US Realm • US Realm Header (V3): Preferred Language recordTarget/patientRole/patient/languageCommunication/languageCode • ISO 639 language code set HL7 CDA R2 Implementation Guide: Public Health Case Report – eICR Release 2, STU Release 3.1 - US Realm • US Realm Header (V3): Preferred Language recordTarget/patientRole/patient/languageCommunication/languageCode • ISO 639 language code set
eCR FHIR	US Public Health Patient: Patient.communication.language • The language which can be used to communicate with the patient about his or her health.
ELR	• PID-15: Primary Language • Vocabulary Standard: ISO 639-2
USCDI v6	Preferred Language • Applicable vocabulary: IETF (Internet Engineering Task Force) Request for Comment (RFC) 5646, "Tags for Identifying Languages", September 2009. Adopted at 45 CFR 170.207(g)(2) • Part of Patient Demographics/Information data class
USCDI+ Public Health: Case Reporting Use Case	Preferred Language • Definition: Documented language that the patient would like to use when discussing health care • Applicable vocabulary: IETF RFC 5646 Language Tags • Part of Patient Demographics/Information data class

Educational Attainment

Collecting data on education allows public health professionals to analyze patterns in health behavior, access to preventive care, and susceptibility to certain health conditions across educational levels. Patient education level not only affects access to preventive care but also may limit comprehension of potentially relevant education materials. Effective patient materials must be written at the appropriate literacy level for all patients to comprehend. The Gravity Project has identified “Less than High School Education” as one of their defined SDOH-related data elements.

The U.S. Census Bureau conducts multiple surveys that collect data on educational attainment, including the American Community Survey (ACS), Current Population Survey (CPS), and CDC’s National Health Interview Survey (NHIS). All three surveys collect the highest degree or level of school that a person has completed, which is distinct from the level of schooling someone is currently attending. The ACS and CPS collect this information at the same level of granularity, while the NHIS has a comparable but different value set (Table 8).

Table 8. Value sets for educational attainment used by the American Community Survey, Current Population Survey, and National Health Interview Survey.

ACS & CPS	NHIS
No schooling completed	Never attended/kindergarten only
Nursery school	
Grades 1 through 11	Grade 1-11
12th grade—no diploma	12th grade, no diploma
Regular high school diploma	High School Graduate
GED or alternative credential	GED or equivalent
Some college credit, but less than 1 year of college	Some college, no degree
1 or more years of college credit, no degree	
Associates degree (for example: AA, AS)	Associate degree: occupational, technical, or vocational program
	Associate degree: academic program
Bachelor’s degree (for example: BA. BS)	Bachelor's degree (Example: BA, AB, BS, BBA)
Master’s degree (for example: MA, MS, MEng, MEd, MSW, MBA)	Master's degree (Example: MA, MS, MEng, MEd, MBA)
Professional degree beyond bachelor’s degree (for example: MD, DDS, DVM, LLB, JD)	Professional School degree (Example: MD, DDS, DVM, JD)
Doctorate degree (for example, PhD, EdD)	Doctoral degree (Example: PhD, EdD)
	Refused
	Don't Know

Two CDC MMGs currently collect data elements related to education (Table 9). The Carbon Monoxide MMG includes a data element called “Education” that indicates the highest degree or level of school completed at the time of the event. The RIBD MMG includes a set of data elements related to education; these include: Is the patient (15-24 years only) currently attending college, their grade in school, and the name of the college or university.

Table 9. Data elements related to education from CDC Message Mapping Guides (MMGs).

MMG	Data Element Name	Identifier	Data Element Description	Value Set Name
Carbon Monoxide	Education	80913-7	Indicate the highest degree or level of school completed at the time of the event.	Education (CO)
RIBD	Attending Higher Education College	224311000	Is patient (15-24 years only) currently attending college?	Yes No Unknown (YNU)
RIBD	Grade in School	64990-5	Year of college	School Year (RIBD)
RIBD	Name of College / University	INV1092	Name of College / University	N/A – Free Text

The DSWG aimed to define a value set for educational attainment that can be utilized by public health for health equity considerations during data analysis and accurately reflect a person's self-assessment of their educational attainment (Table 10). The value sets for national surveys outlined in Table 8 are somewhat collapsed in the value set recommended here. Additionally, the value sets for the national surveys do not include a specific category for a technical/vocational school certification, although this represents a common educational attainment separate from the traditional college/university-based educational pathway. While inclusion of "Technical/Vocational certification" in the value set for Educational Attainment is misaligned with these national survey data sources, the DSWG supports its inclusion here as there may be health-related outcome differences between this population and those with a high school diploma.

[page intentionally left blank]

Table 10. Core Surveillance Data Element for Routine CDC Case Notification – Educational Attainment

Educational Attainment	
Definition	Highest degree or level of school completed by the person at the first epidemiologically relevant date associated with the case (i.e., CCCD*)
Standard Code for Transmission to CDC	LOINC: 82589-3 Highest level of education
Value set or coding system	• Less than 9th grade • 9th-12th grade, no diploma • High school graduate (includes equivalency, i.e., GED) • Some college/university, no degree • Technical/vocational certification • Associate's degree • Bachelor's degree • Graduate or professional degree • Unknown
Abstracted/ Derived/ Created	Abstracted
Recommendations for Input Data Source(s)	• Case investigation • eCR
Notes	The American Community Survey does not include technical or vocational certification within the data element for educational attainment. This difference in data collection should be considered if utilizing ACS data as a comparison or denominator during analysis. This value set is based on the educational system within the United States; the most appropriate equivalent selection should be made for people whose educational setting is/was outside of the U.S.
Data Source Details **	
eCR CDA	HL7 CDA R2 Implementation Guide: Public Health Case Report, Release 2 STU Release 1.1 - US Realm • A standardized method of transmitting educational attainment is not outlined, but relevant information may be found as a Social History Observation within the Social History Section using Social History Type of 105421008 | Educational achievement (observable entity). • ClinicalDocument/component/Social History Section/Social History Observation/observation/value HL7 CDA R2 Implementation Guide: Public Health Case Report – eICR Release 2, STU Release 3.1 - US Realm • A standardized method of transmitting educational attainment is not outlined, but relevant information may be found as a Social History Observation within the Social History Section using Social History Type of 105421008 | Educational achievement (observable entity). • ClinicalDocument/component/Social History Section/Social History Observation/observation/value
eCR FHIR	N/A
ELR	N/A
USCDI v6	No direct comparison to this data element exists, but relevant information may be found in the following:

	SDOH Assessment • Definition: Screening questionnaire-based, structured evaluation for a Social Determinants of Health-related risk. Examples include but are not limited to food, housing, and transportation security. • Applicable vocabulary: LOINC, SNOMED CT • Part of Health Status Assessments data class
USCDI+ Public Health: Case Reporting Use Case	N/A

* The Calculated Case Counting Date (CCCD) is the earliest date available of six component dates and is referenced as the earliest epidemiologically relevant date associated with the case. Data Standardization: Workgroup Proposal for Standardized Case Counting Date

** Additional information on input data sources can be found in Appendix A.

[page intentionally left blank]

Insurance

Health insurance is a social factor that directly impacts access to healthcare and therefore a person's health. Different types of health insurance (e.g., private insurance, Medicare, Medicaid, uninsured) are often linked to various socioeconomic factors, making insurance coverage a useful proxy for social determinants of health. The Gravity Project defines a data element of "Health Insurance Coverage Status" as the "Documentation of presence and type of insurance". By analyzing insurance data, public health entities can identify disparities and target interventions to promote health equity.

The U.S. Census Bureau broadly defines health insurance coverage to include both public and private payers that finance medical expenditures for a specific population across diverse settings. The ACS, CPS, the Survey of Income and Program Participation (SIPP), and NHIS all assess health insurance coverage. An example of a question from the ACS is "Are you currently covered by any of the following types of health insurance or health coverage plans?" with the following list of responses:

- Insurance through a current or former employer or union (of yours or another family member)
- Insurance purchased directly from an insurance company (by you or another family member)
- Medicare, for people 65 and older, or people with certain disabilities
- Medicaid, Medical Assistance, or any kind of government-assistance plan for those with low incomes or a disability
- TRICARE or other military health care
- VA (enrolled for VA health care)
- Indian Health Service
- Any other type of health insurance or health coverage plan – Specify

The RIBD MMG is the only condition-specific MMG that includes an "Insurance" data element. This data element is a repeating element, meaning it takes a "select all that apply" approach. The value set contains the following values:

- Incarcerated
- Indian Health Service
- Managed Care (Private)
- Managed Care Unspecified
- Medicaid
- Medicare
- Military/VA insurance
- Other
- Private Health Insurance
- Uninsured
- Unknown

[page intentionally left blank]

Table 11. Core Surveillance Data Element for Routine CDC Case Notification – Healthcare Insurance or Access Type

Healthcare Insurance or Access Type	
Definition	Person's type of health insurance coverage or other healthcare access, select all that apply
Standard Code for Transmission to CDC	TBD
Value set or coding system	All options that apply may be selected from the following value set: Private insurance (e.g., from employer, from Marketplace) Medicare Medicaid TRICARE or other military-associated health care Veterans' Health Administration health care Indian Health Service funded (includes Tribe operated or Urban Indian health center) Correctional/Detention facility healthcare service Other public insurance (e.g., CHIP, community health care plan sponsored by local government) Other Uninsured/Self-pay Unknown
Abstracted/ Derived/ Created	Abstracted
Recommendations for Input Data Source(s)	• Case investigation • eCR
Notes	There was weak support for inclusion of this data element in the core data for routine case notification to CDC.
Data Source Details*	
eCR CDA	N/A
eCR FHIR	N/A
ELR	N/A
USCDI v6	Coverage Type • Definition: Category of healthcare payers, insurance products, or benefits. Examples include but are not limited to Medicaid, commercial, HMO, Medicare Part D, and dental. • Applicable vocabulary: N/A • Part of Health Insurance Information data class
USCDI+ Public Health: Case Reporting Use Case	N/A

*Additional information on input data sources can be found in Appendix A.

Appendix A. Input Data Standards and Data Sources for SDOH Data

See Tables 4, 5, 7, 10 and 11 for a summary of where information related to the SDOH core data elements may be found within data transmitted to public health using existing standards (eCR, ELR) or how data may be captured within EHR systems for electronic data exchange (USCDI).

Using Patient Address to Map Residence Type

Both eCR and ELR contain fields to transmit a patient address, and USCDI version 6 defines “Current Address” as “Place where a person is located or may be contacted". A patient address may contain information that is relevant to residence type. For instance, if the address listed for a patient is the address of a known correctional facility or skilled nursing facility, this conveys they may live in that kind of congregate living facility. In some cases the address field might include a term like ‘homeless’. An address may also show as ‘no fixed address’ or a general delivery address at a post office, which are not necessarily indicators of homelessness but can be indicators that the person does not have a permanent residence and thus further investigation into the person’s living situation is warranted.

USCDI and USCDI+

The most recently finalized version of USCDI (version 6) includes four SDOH specific data elements, although the descriptions of these data elements are broad and therefore do not neatly align with the data elements within this document:

- SDOH Assessment, which is defined as “Screening questionnaire-based, structured evaluation for a Social Determinants of Health-related risk.” This data element is currently included in the Health Status Assessments data class, which includes various health-related assessments such as functional status, disability status, and use of alcohol or other substances.
- SDOH Problems/Health Concerns, which is defined as “Social Determinants of Health-related concerns, conditions, or diagnoses.” This data element is currently included in the Problems data class, which is described as “condition, diagnosis, or reason for seeking medical attention” and contains a separate “Problems” data element.
- SDOH Goals, which is defined as “Desired future state for an identified Social Determinants of Health-related concern, condition, or diagnosis.” This data element is currently included in the Goals and Preferences data class.
- SDOH Interventions, which is described as “Actions or services to address an identified Social Determinants of Health-related concern”. This data element is currently included in the Procedures data class.

Preferred Language and Healthcare Insurance or Access Type have data elements present in USCDI v6 that support the data elements defined by the DSWG (see Tables 7 & 11). Additional relevant data elements for Housing Status, Primary Living Situation, and Educational Attainment are not yet within a published version of USCDI but exist at “Level 0”, which means they are not represented by current terminology standards, only captured in limited settings or exchanged in limited environments, and the use cases apply to a limited number of care settings (see Table 12).

Table 12. Additional Relevant SDOH Data Elements in USCDI at Level 0.

Data Class	Data Element	Definition
Social Determinants of Health	Education Level	Highest level of education obtained.
Social Determinants of Health	Congregate Living	Shared housing includes a broad range of settings, such as apartments, condominiums, student or faculty housing, national and state park staff housing, transitional housing, and domestic violence and abuse shelters.
Social Determinants of Health	Housing Instability and Homelessness	Currently consistently housed, but experiencing any of the following circumstances in the past 12 months: being behind on rent or mortgage, multiple moves, homelessness; or

		currently living in a shelter, motel, temporary or transitional living situation, scattered site housing, or not having a consistent place to sleep at night; or lacking a fixed, regular, and adequate nighttime residence.
Social Determinants of Health	Incarceration History	In the past year, have you spent more than 2 nights in a row in a jail, prison, detention center, or juvenile correctional facility?

USCDI+ supports domain specific data element lists and includes a set of data classes for Public Health use, including data classes for a specific use case of Case Reporting. Relevant data elements within the Case Reporting use case are contained within the table for each core data element (Tables 4, 5, 7, 10 and 11). While Educational Attainment and Insurance Coverage type do not have relevant data elements within the Case Reporting use case, other use cases within USCDI+ have relevant data elements (Table 13).

Table 13. Additional Relevant SDOH Data Elements in USCDI+ outside of the Case Reporting use case.

Domain - Use Case	Data Class	Data Element	Definition
Behavioral Health – Comprehensive Care	Health Insurance Information	Educational Attainment	May refer to current grade for children, or highest education level for adults (e.g. high school or general equivalency).
• Behavioral Health – Comprehensive Care • Quality – Quality Overarching & Quality Draft v1 • Maternal Health – Maternal Health Overarching	Health Insurance Information	Coverage Type	Category of health care payers, insurance products, or benefits. Examples include but are not limited to Medicaid, commercial, HMO, Medicare Part D, and dental.
• Behavioral Health – Comprehensive Care • Quality – Quality Overarching & Quality Draft v1 • Maternal Health – Maternal Health Overarching	Health Insurance Information	Coverage Status	Presence or absence of health care insurance.