Prepared for:

Centers for Disease Control and Prevention

Centers for Medicare & Medicaid Services Alliance to Modernize Healthcare (Health FFRDC) A Federally Funded Research and Development Center

Clinical and Community Health Data Initiative

Contract No. 75FCMC18D0047

Task Order No. 75D30120F09743

CODI Data Models Implementation Guide

For the North Carolina Site (2021–2022)

Version 4.2.1

January 31, 2022

The views, opinions, and/or findings contained in this report are those of The MITRE Corporation and should not be construed as official government position, policy, or decision unless so designated by other documentation. This guide may serve as a reference or framework for others implementing similar clinical and community health data solutions. However, the CODI Implementation Guide does not represent official views or guidance of CDC.

Version 3.0 Approved for Public Release: 20-0932. Distribution Unlimited.

Version 4.1 Approved for Public Release: 21-3933. Distribution Unlimited.

Record of Changes

Version	Date	Author / Owner	Description of Change
1.0	Aug 7, 2019	P. Mork / Health FFRDC	Version approved for public release
1.1	Sep 4, 2019	P. Mork / Health FFRDC	Updated cover page to include CDC disclaimer
1.2	Oct 14, 2019	P. Mork / Health FFRDC	Additional clarifications based on feedback from CDC
2.0	Nov 6, 2019	P. Mork / Health FFRDC	1) Added the BMIAGE reference table from CDC, which was subsequently deleted. 2) Changed PROGRAM_MODE to SESSION_MODE to account for programs with multiple modes. 3) Synchronized the CODI Research Data Model with PCORnet version 5.1 by a) updating the codes for VITAL_SOURCE_TYPE and b) changing the datatype for FACILITY_LOCATION within an ENCOUNTER.
3.0	Mar 30, 2020	P. Mork / Health FFRDC	1) Removed the BMI data elements from VITAL and the BMIAGE table. 2) Added three new data elements to the PROGRAM table, which are documented in section 5.8. 3) Added the CURRICULUM_ COMPONENT table, which is documented in section 5.3.
4.0	Sep 08, 2021	C. Macheret / Health FFRDC	1) Removed CODI IDENTIFIER and IDENTITY_HASH_BUNDLE tables from the RLDM and added PRIVATE_DEMOGRAPHIC and PRIVATE_ADDRESS_HISTORY tables from PCORnet CDM 6.0. 2) Added IMMUNIZATION, CONDITION, and PRO_CM from PCORnet CDM 6.0 to the RDM. 3) Added SDOH_EVIDENCE_INDICATOR and HOUSEHOLD_LINK to the RDM. 4) Expanded definitions in the RDM ancillary tables that were previously constrained to children to include adults as well as definitions previously constrained to obesity to all chronic disease and chronic comorbidities.
4.1	Dec 2, 2021	K. Mork / Health FFRDC	Added ENROLLMENT and two tables related to PREGNANCY.
4.1.1	Jan 25, 2022	K. Mork / Health FFRDC	Clarified that pregnancy tables are relevant to any data owner that can populate the data. Clarified that WIC is not a CBO. Revised language for SDOH to align with CDC guidance. Added examples of how to populate SDOH data and evidence indicators.
4.1.2	Feb 7, 2022	K. Mork / Health FFRDC	Added section 3.5 to provide guidance to clinical organizations, CBOs, and government benefits organizations.

4.1.3	April 01, 2022	C. Macheret / Health FFRDC	1) Replaced “Data Partner” with “CODI Implementer” or “Implementer” where appropriate to avoid having to distinguish between data owners implementing CODI for their own data versus data partners implementing CODI for another data owner. 2) Clarified guidance on selection of private address data for the record linkage agent. 3) Corrected cardinality in the RLDM for the relationship between DEMOGRAPHIC and the private tables because CODI PPRL requires only one record per individual. 4) Added version numbering information to section 5.1 describing change control. 5) Corrected the data dictionary appendix by removing duplicate rows for the following tables: ALERT, CENSUS_DEMOG, and adding missing rows in the following tables: LINK and HOUSEHOLD_LINK 6) Added the record ids in the LINK table and HOUSEHOLD_LINK table. The former has record id equal to LINKID returned from linkage agent. The second table has meaningless record id just for purposes of following record id data modeling convention. Changes here match the DDL in CODI GitHub.
4.1.4	June 1, 2022	C. Macheret / Health FFRDC	1) Alex Beede improved readability 2) Updated definitions and CODI overview in Section 2.1 3) Removed duplicate rows from tables in the data dictionary appendix 4) Added detail to Section 4.1.5
4.1.5	September 8, 2022	C. Macheret / Health FFRDC	1) Removed duplicated rows in data dictionary appendix 2) Included additional fields to CODI PRIVATE_ADDRESS_HISTORY from PCORnet to accommodate multiple addresses for use in geocoding census location history if available for an individual. Updated associated guidance in section 4.1.5 3) Added GENDER_IDENTITY to the DEMOGRAPHIC table. Added GENDER_IDENTITY_TYPE to the code tables. An explanation was added to section 4.1. 4) Changed data type of all ID fields in the COST table from integer to ID (a string) to permit FKs for CODI tables whose IDs are strings. 5) Updated the CENSUS_DEMOG table field type for CENSUS_DATA_SRC to char(26) from char(16) 6) Fixed underscore in attribute names: PAT_MIDDLE_NAME, PROVIDERID
4.2 draft for review	October, 2022	C. Macheret / Health FFRDC	1) Removed CENUS_DEMOG: Removed foreign keys to same table from PROGRAM and CENSUS_LOCATION table. The latter should not be constrained as it should allow geocodes even at the county or state level. Going forward, the DCC will provide the population statistics for census locations as needed. This is the decision of the CODI ISG.

			2) Updated link to HL7 Null Flavors from previous FHIR version 3. Fixed additional underscore inconsistencies in field names. 4) Changed ENROLLMENT table name to PROGRAM_ENROLLMENT due to naming conflict with PCORnet CDM. Also changed PK to match. 5) Explained SDOH data source scenarios and associated CODI mapping for more explicit guidance on inserting SDOH indicator records. 6) Moved information about the Gravity Project and CODI SDOH category alignment to an appendix. 7) Removed HOUSEHOLD_LINK_ID field from HOUSEHOLD_LINK to match pattern used for LINK table (which is different than the single-field technical primary key pattern for most of the CODI data model). The HOUSEHOLD_LINK table PK is now a composite key, as with the LINK table. 8) Added missing SDOH Category Code: SOCIAL_CONNECTION_DOMAIN 9) Added a sub-section in North Carolina specific guidance appendix explaining the mapping of the CODI DM IG data types to Postgres SQL and SAS data types. 10) Changed data types for the latitude and longitude fields in CENSUS_LOCATION and PROGRAM.
4.2 Released	January 06, 2023	C. Macheret / Health FFRDC	1) Added 'PH' for 'Permanent Housing' as a new Asset Type to accommodate NCCEH project data. 2) Updated definition of HOUSEHOLDID 3) Included PCORnet CDM table, HARVEST to collect CODI ETL refresh information. 4) Updated CODI concept overview diagram (figure 1) to remove reference to population demographics which is no longer part of the CODI data model (see previous update note regarding CENSUS_DEMOG). 5) Added rules for data fields missing or null in source system, but mandatory in target system, to the General Guidance section. 6) Corrected constraint on ADDRESS_USE field in PRIVATE_ADDRESS_HISTORY 7) Added guidance for LOCATION_ADDRESS of programs that are designed and administered online or at home. In these cases, LOCATION_ADDRESS should say 'Virtual'. 8) Updated the table listing primary data types for named value sets in the physical implementation section of Appendix B. 9) Made minor corrections in documentation in the appendices C and D. 10) Making available a spreadsheet version of the appendices in C and D.

			11) Converted SDOH_CATEGORY codes from long names to two characters and moved long name into the definition.
4.2.1	January 30, 2023	C. Macheret / Health FFRDC	1) Added DESTINATION_PROGRAMID to REFERAL as an association property. Removed mandatory constraint on source and destination ORGANIZATION. Updated REFERRAL definition to allow either clinical provider or program as the referral destination.

Table of Contents

List of Figures

List of Tables

1. Introduction

The Centers for Disease Control and Prevention (CDC) promote health; prevent disease, injury, and disability; and prepare the nation for emerging health threats. The CDC's Division of Nutrition, Physical Activity, and Obesity and the Center for Surveillance, Epidemiology, and Laboratory Services partnered with the Centers for Medicare & Medicaid Services (CMS) Alliance to Modernize Healthcare federally funded research and development center (Health FFRDC), as well as local clinical and health partners, for the Clinical and Community Data Initiative (CODI). CODI expands the ability to capture, standardize, integrate, and query existing patient-level electronic health records (EHRs) and community-based program data via a common data model.

The CODI data models are an extension of the Patient Centered Outcomes Research Network's (PCORnet) Common Data Model (CDM),[1] a standard data model representing anonymized patient-level data for research. CODI augments the PCORnet CDM with a representation of chronic disease prevention programs and individual-level program participation data. Researchers can use CODI to access anonymized longitudinal record data for individuals that includes both clinical history and local public health program data.

This document describes how CODI implementers and end-users should interpret the CODI data tables. It also provides best practices for addressing situations in which CODI implementers may identify multiple ways to populate the CODI data models with local data. CODI end-users can query across the health and intervention data shared by multiple data owners.

1.1 Background

Version 4 of the CODI Data Models Implementation Guide (DM IG) is an update to the CODI DM IG, Version 3.0 (March 2020) that was developed for the CODI participants in Colorado (CODI@CO) starting in 2019. Version 4 accommodates expanded scope requests from the Scoping and Use Case Subgroup of the CODI Collaborative Work Group (CCWG) tailored to the NC implementation of CODI, (CODI@NC). CODI@NC, starting in January 2022, began implementation using Version 4.1 of the DM IG, which has driven additional changes reflected in Version 4.2 and may continue to drive subsequent revisions to the IG as implementation progresses.

To determine the scope of available, relevant data owned by North Carolina organizations, the CCWG and its partners performed a clinical community linkages assessment (CCLA) and a technical environmental scan (TES) of North Carolina healthcare delivery and community-based organizations who may own data relevant for CODI@NC.

The Health FFRDC used the CCLA TES findings, the requested scope expansion, and the existing Version 3.0 data models as inputs to develop an analysis and subsequent data model change recommendations documented in the CODI@NC Gaps Analysis report.[2] The CODI data models and implementation guidance in Version 4 are a culmination of that analysis and

[1] The acronym CDM can refer to common data models from different authorities. Within this document, when used without qualification, it refers to the PCORnet Common Data Model.

[2] Contact Division of Nutrition, Physical Activity, and Obesity at CDC.gov for access to this report.

feedback from CCWG. Clarifications to the implementation guidance that are not changes to the data model are expected during the North Carolina implementation process resulting in subsequent updates of the CODI Implementation Guide. For the IG change control process, see Section 5.1.

The CODI data models comprise two distinct data schemas: 1) the CODI Research Data Model (RDM), representing data needed to answer CODI end user queries on the health status, health intervention participation, and community-based program participation of individuals; and 2) the CODI Record Linkage Data Model (RLDM), needed for matching an individual's records across different data owners and for matching individuals to a household.

CODI's data models incorporate parts of the PCORnet CDM and introduce ancillary tables either borrowed from other data models or designed specifically for new functionality to a PCORnet or PCORnet-compatible clinical data network. Two of CODI's ancillary tables are adopted from the Colorado Health Observation Regional Data Service (CHORDS) virtual data warehouse (VDW) and the Observational Medical Outcomes Partnership (OMOP) common data model because they satisfied CODI data requirements. The CODI project designed the remaining ancillary tables for CODI functionality not already supported by PCORnet, CHORDS VDW, or OMOP at the time CODI DM was last updated.

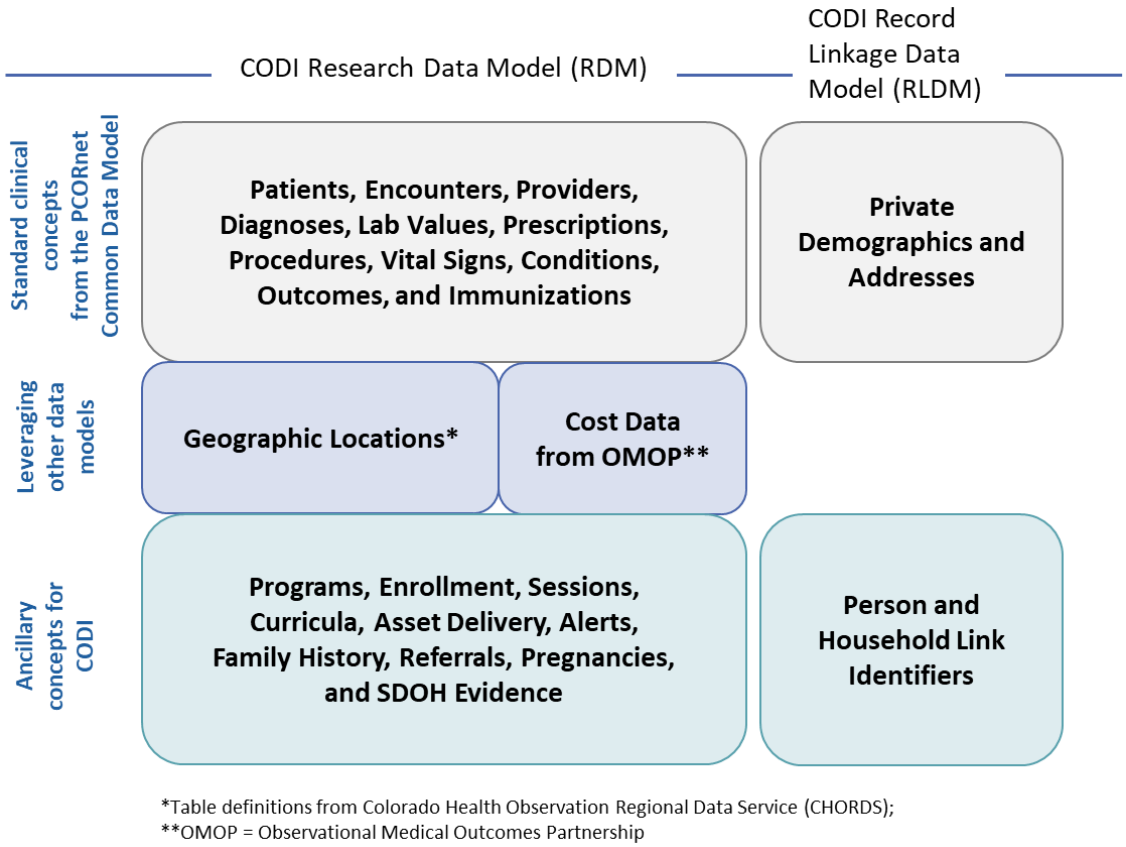


Figure 1. Overview of CODI Research and Record Linkage Data Models

Figure 1 illustrates the major components of the CODI data models and the provenance for those components.

1.2 Purpose

The purpose of this document is to provide the guidance necessary for implementers to build and populate the CODI data models. Toward that end, this document provides:

- Descriptions of the two CODI data models: RDM and RLDM
- General expectations regarding how to populate the tables in the CODI data models
- Specific guidance regarding individual data elements of the CODI data models

The data dictionary appendices to this guide are useful to both implementers and end-users for understanding the meaning of the data elements in CODI data models. Some of the content in the appendices reflect dictionary content from models that CODI has adopted (e.g., PCORnet CDM), and therefore are only contained in this document as a convenience to the reader, not as the official source of record. The adopted content is marked as such.

1.3 Scope

This document comprises the general implementation guidance for the ancillary tables (i.e., the lowest layer in Figure 1) and CODI supplemental guidance necessary for CODI's specific use of adopted tables.

This document is not the authoritative source for PCORnet CDM and the other, incorporated data models (the top two layers in Figure 1). Detailed implementation guidance for those can be found in the following documents:

- PCORnet Common Data Model v.6.0 Specification [3]
- CHORDS VDW 3.5 Data Model Manual [4]
- OMOP Common Data Model v.6.0 Specifications [5]

In addition to implementation guidance, this document provides the data definitions for all tables in the CODI data models, both incorporated and ancillary.

Should any of the adopted data models undergo revision, the CCWG must assess the impact to CODI and choose whether to continue with the model versions already incorporated in the current CODI version, or to appropriately revise and re-release the CODI IG.

Data owners or partners are encouraged to supplement the CODI implementation guide with their own data mapping and implementation specifications unique to their distinct information systems. Those data owner guides provide refinements and additional information but shall not contradict information presented here.

Because implementation of the full RDM may be initially complex, this guide recommends an implementation priority in Section 2.2.2 for RDM tables. Implementers familiar with Version 3

[3] https://pcornet.org/wp-content/uploads/2020/12/PCORnet-Common-Data-Model-v60-2020_10_221.pdf

[4] The Colorado Health Observation Regional Data Service (CHORDS) virtual data warehouse (VDW) document is available by sending a request to CODI@cdc.gov.

[5] https://github.com/OHDSI/CommonDataModel/wiki

of the CODI Data Models should take note of the removal of components of RLDM as documented in Section 2.2.1.

1.4 Audience

The primary audience for this document is the technical staff of the CODI implementers—those individuals directly responsible for populating the CODI Data Models using data from participating data owners. The secondary audience includes project staff indirectly responsible for implementation and for potential new CODI participants trying to assess the feasibility of implementing the CODI Data Models. Chronic disease researchers are likely most interested in the data model documentation appearing in the appendices or as a stand-alone data dictionary, needed to formulate data queries.

1.5 Document Organization

This document is organized as follows:

Table 1: Document Organization

Section		Purpose
Section 2:	CODI Overview	Provides an overview of the CODI data models and introduces common definitions
Section 3:	General Guidance	Provides general guidance for implementers
Section 4:	Specific Guidance	Provides guidance on how to implement the changes made by CODI to existing PCORnet CDM tables and how to implement the CODI ancillary tables
Section 5:	Additional Resources	Provides additional information on contacts and governance for this implementation guide
Appendix A:	CODI SDOH Categories and The Gravity Project	Explanation of Social Determinants of Health Domains and alignment with The Gravity Project
Appendix B:	Additional Guidance for CODI@NC	Provides specific guidance for implementation in North Carolina
Appendix C:	RLDM Data Definitions	Presents the detailed RLDM data dictionary
Appendix D:	RDM Data Definitions	Presents the RDM data dictionary
Appendix E:	CODI Research Data Model Codeset Details	Lists of value sets
Acronym List		Defines the acronyms used in this document
List of References		Lists the sources used in preparing this document

2. CODI Overview

This section begins with a brief description of CODI's concept of operation and defines the CODI roles that are relevant to implementing the CODI Data Models. This is followed by an overview of the CODI Data Models.

2.1 CODI Operational Concept and Roles

Different organizations within a community collect different types of data on an individual's health or health behavior. CODI links the data systems of these organizations to build individual-level anonymized, longitudinal health records. CODI users then query the CODI system to access health and health behavior data relevant to their research or program questions.

In the CODI model, the organizations that own the data of interest are called **data owners.** Data owners may be:

- Clinical healthcare providers, such as hospital systems, community health centers, or individual providers
- Community-based organizations (CBO)
- Government organizations such as Public Health Departments
- Other organizations that collect health or social risk factor data

Data owners contribute to CODI by allowing their data to be linked with that of other data owners within the same **clinical-community distributed data network** through a process called Privacy Preserving Record Linkage (PPRL), which is performed by the **linkage agent**.

The linkage agent is an organization that links data on behalf of CODI implementers. The linkage agent receives encrypted personally identifiable information (PII) from data owners and produces unique LINKIDs, which can link an individual's data across organizations. Figure 2 shows a linkage agent receiving hashed data from two data owners and linking individuals across those data owners. In practice, the linkage agent will perform this linkage across all data owners within the CODI clinical-community distributed data network at regular intervals.

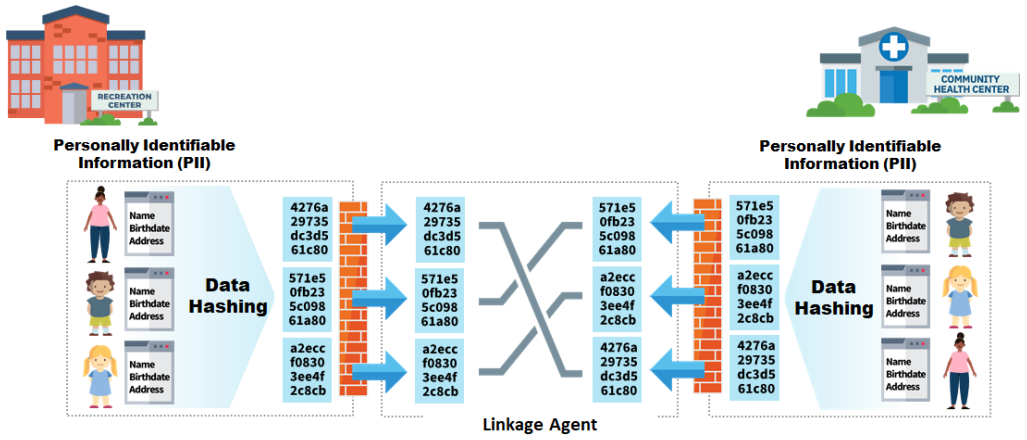


Figure 2. Example of Privacy Preserving Record Linkage Performed by a Linkage Agent

To facilitate the linkage, data owners must map their data (including the LINKIDs) to the CODI model so the data can be queried and assembled into a longitudinal record in a standardized way by the **data coordinating center**. Figure 3 shows how CODI users (e.g., researchers, community-based program evaluators) interact with the data coordinating center, which distributes the research queries throughout the CODI network. The data coordinating center assembles the results into longitudinal records, which are then sent to the CODI users.

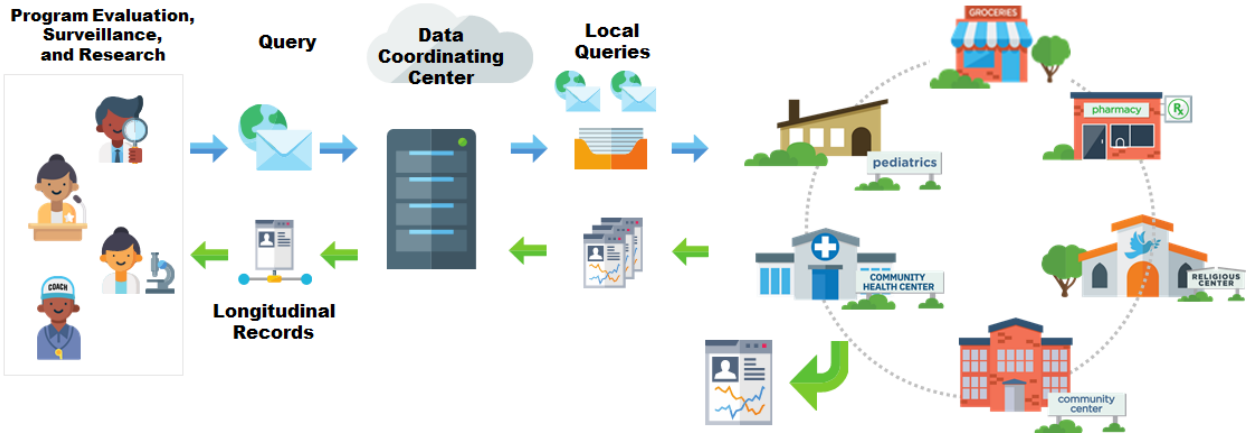


Figure 3. A clinical-community distributed data network

If a data owner does not have the resources for these tasks, they may rely on an intermediary **data partner** to help with the data mapping and host their data. For the purposes of this implementation guide, any organization who is responsible for populating the CODI data model is referred to as a **CODI implementer,** whether they are a data owner or a data partner acting on behalf of a data owner. This implementation guide is intended to assist CODI implementers with populating the CODI data model.

2.2 CODI Data Models

The next two subsections provide a conceptual overview of the RLDM and RDM.

2.2.1 Record Linkage Data Model

CODI uses record linkage to construct a longitudinal record of an individual's health from the information supplied by multiple data partners, while protecting anonymity. CODI uses the individual-to-household linking to enable analysts to explore correlations among household members in their behavior and health.

The RLDM defines the data tables and data elements needed to perform the privacy-preserving record linkage process. It includes two tables from PCORnet CDM designed to contain unencrypted PII. Those two tables must be implemented by a data owner or partner in a secure repository, separated from health and behavior data, as explained in the CODI Privacy Preserving Record Linkage Implementation Guide.[6] Table 2 summarizes the conceptual components of the RLDM. Appendix C contains the RLDM data dictionary with table and attribute definitions.

Table 2. Conceptual Components of the RLDM

Concept	Table	Description
Private individual demographics	PRIVATE_DEMOGRAPHIC	Includes PII on an individual that is not shared but is used to create anonymous identifiers
Private individual addresses	PRIVATE_ADDRESS_HISTORY	Includes personal address information for an individual that is not shared but is used to create anonymous household identifiers and to geocode an individual's location
Anonymous identifiers	LINK HOUSEHOLD_LINK	Includes the anonymous identifiers used to link information on individuals and households back to the DEMOGRAPHIC table

The RLDM is revised considerably from the model defined in the CODI DM IG, Version 3.0 for Colorado. The PCORnet CDM private tables replace the IDENTIFIER table. Updates to the CDM, particularly the addition of the private tables, render the IDENTIFIER table obsolete.

[6] https://github.com/mitre/codi/blob/main/CODI%20PPRL%20Implementation%20Guide.pdf

2.2.2 Research Data Model

The RDM provides the data tables and data elements needed to answer selected chronic disease research and program evaluation questions. Table 3 summarizes the major conceptual components of the RDM. High-priority data tables are required for a minimum CODI implementation. Medium-priority tables should only be implemented by data partners with the resources and local interest to do so. Low priority tables are included for completeness—there are no plans to implement them at this time.

Table 3. Conceptual Components of the RDM

Concept	Table	Description
The individual	DEMOGRAPHIC	Includes demographic information about a patient or program participant
Family health risk factors	FAMILY_HISTORY	Includes details about any family members' health conditions
Clinical care	ENCOUNTER DIAGNOSIS LAB_RESULT_CM PRESCRIPTION PROCEDURE PROVIDER VITAL	Includes information about an individual's interactions with the healthcare delivery system
Clinical or self-reported COVID-19 status	IMMUNIZATION CONDITION LAB_RESULT_CM	Includes information on an individual's COVID-19 disease and vaccination status
Individual-level social determinants of health	CONDITION DIAGNOSIS PRO_CM SDOH_EVIDENCE_INDICATOR	Includes responses to questionnaires for collecting an individual's social determinants of health and a person-specific map (SDOH_EVIDENCE_INDICATOR) of where to find that information in the RDM
Chronic disease-related interventions	PROGRAM CURRICULUM_ COMPONENT PROGRAM_ENROLLMENT SESSION ASSET_DELIVERY	Includes details about intervention aims and settings (PROGRAM); how the interventions are structured (CURRICULUM COMPONENT); who is enrolled in programs and who is administering the intervention and how (PROGRAM_ENROLLMENT, SESSION); and if an asset (e.g., food, money) was provided (ASSET DELIVERY)
Pregnancy	PREGNANCY PREGNANCY_OUTCOME	Includes information about prenatal care, delivery, and postnatal circumstances
Referrals	REFERRAL	Includes incoming and outgoing referrals within and across organizations

Concept	Table	Description
Clinical decision support	ALERT SESSION_ALERT	Includes details about the types of clinical alerts (ALERT) and when they triggered (SESSION ALERT)
Cost of care	COST	Includes information about the amounts charged
Location area of Individual	CENSUS_LOCATION	Census location links individuals to geographic areas defined by the Census, based on their current and past known home addresses. A geographic area shall not be more specific than a Census Tract (an area bigger than a block group, but usually smaller than a county). Census locations can be cross-referenced with Census data[7] to obtain community demographic context for individuals.

[7] Specifically, from the American Community Survey

3. General Guidance

This section provides general guidance for CODI implementers to populate the CODI data models. This guidance applies to the tables in the RDM and RLDM.

3.1 Data Cleaning Expectations

In general, the CODI Data Models should be populated with structured data extracted from the EHR or other information technology (IT) systems, unless specified otherwise. For example, CODI implementers will not populate a data element in the RDM based on an analysis of free text, such as a progress note, nor should they attempt to suppress implausible values. CODI end-users can perform any data cleaning or inferencing post-hoc based on their data needs.

Exceptions to this general guidance are listed for specific data elements. For example, an exception is made for the process-related data elements of the SESSION table. This exception exists because only the data owners and their implementors can determine which process steps necessarily follow from observations recorded in their systems.

3.2 Data Transformation Expectations

3.2.1 Mapping Codesets

Mapping from a local codeset to a CODI codeset is anticipated and does not constitute data cleaning. CODI implementers **should** map a data owner's native codes to CODI.

3.2.2 Missing Data

PCORnet CDM uses the Health Level 7 International (HL7) conventions [8] for missing or unknown values; these rules will therefore apply to the CODI ancillary tables as well:

- A data field that is not present in the source system uses a null value. See exception below, for mandatory fields in CODI.
- If the source value is null or blank, PCORnet CDM uses NI (no information). See exception below, for mandatory uncoded text fields in CODI.
- If the source value is an explicit unknown value, PCORnet CDM uses UN (unknown).
- When the source value cannot be mapped to PCORnet CDM, CDM uses OT (other).

Certain mandatory data fields common in clinical data models might not be common in community-based organization data systems. Because CODI supports both clinical and community-based organization data, certain additional rules for missing data are needed in CODI.

- For a data field that is not present in the source data model but is mandatory and constrained by a value set in the target data model (CODI or PCORnet), use 'NI' in the target data field.

[8] https://terminology.hl7.org/1.0.0/CodeSystem-v3-NullFlavor.html

- For a null or blank data value in the source data field corresponding to a mandatory (uncoded) text data field in the target data model, use a one space character string.

3.2.3 Date Formatting

CODI guidance on dates comes from PCORnet CDM implementation guidance.

An excerpt from the PCORnet CDM version 6.0 implementation guidance follows.

> Because the PCORnet CDM is intended to support multiple Relational Database Management Systems (RDBMS), date format consistency is an issue, given that most RDBMS's have platform-specific native date representation.
>
> To address this issue, each RDBMS will be expected to implement its own native date data type for dates, which will be supported by the Entity Framework technology stack[9]. The CDM will always separate date fields and time fields for consistency and employ a naming convention of suffix "_DATE" or "_TIME".
>
> All times should be recorded within the local time zone. A uniform time stamp or GMT offset is not expected.

3.3 Reference Tables

There are two tables likely to not be populated from an EHR or another IT system. These are ALERT and PROGRAM.

The PROGRAM and ALERT tables must be populated manually. Implementers are encouraged to populate these tables with explicit program and alert values as part of the extract–transform–load (ETL) process that populates the remainder of the RDM. Implementers are further encouraged to test referential integrity to ensure the primary keys for these tables connect properly with the tables that reference them (such as PROGRAM_ENROLLMENT and SESSION).

3.4 Start Date

Each clinical-community distributed data network should establish a **start date** for the data extraction in that network. The start date represents the earliest possible date for which data partners can reliably populate the CODI Data Models. Data partners should use the start date as the earliest event data to populate CODI tables. The purpose of a single start date is for the data from different partners in the same data-sharing network to be comparable. The CODI implementation subgroup of the clinical community network will determine the start date. See Appendix B for any CODI network-specific implementation decisions.

3.5 Guidance by Data-source Organization Type

In the following subsections, we provide guidance for various kinds of organizations including clinical organizations, government benefits organizations, and CBOs. Regardless of organization type, the tables in the RLDM are relevant and required, as they are necessary to participate in PPRL linking individual and household member data across data sources.

[9] https://msdn.microsoft.com/en-us/data/ef.aspx

3.5.1 Clinical Organizations

Clinical organizations include any organization that provides clinical services. In CODI@NC, clinical organizations include Duke, University of North Carolina, and the Durham Local Health Department.

These data partners should create all RDM tables in their data warehouses and populate those tables to the best of their ability. If a table cannot be implemented, the data partner should create the table and leave it empty so queries that reference those tables do not fail.

- High Priority
  - CENSUS_LOCATION
  - DEMOGRAPHIC
  - DIAGNOSIS
  - ENCOUNTER
  - ENROLLMENT
  - LAB_RESULT_CM
  - PREGNANCY
  - PREGNANCY_OUTCOME
  - PRESCRIBING
  - PROCEDURES
  - PROGRAM
  - PROVIDER
  - PRO_CM
  - SDOH_EVIDENCE_INDICATOR
  - SESSION
  - VITAL
- Low Priority
  - ALERT
  - ASSET_DELIVERY
  - CONDITION
  - COST
  - CURRICULUM_COMPONENT
  - FAMILY_HISTORY
  - IMMUNIZATION
  - REFERRAL
  - SESSION_ALERT

3.5.2 Community-Based Organizations

CBOs are organizations in which non-clinical services or assets are delivered. These organizations do not deliver healthcare as would a clinic, hospital, health center, or other organization that provides clinical care. CBO examples include the local YMCA, local

foodbanks, and Divisions of Parks and Recreation. The tables listed as high priority are required for CODI's primary functions, while tables listed as lower priority may be relevant but are optional.

Within the RDM, only the subset of tables listed below are typically relevant[10] to CBOs. CBOs should create all RDM tables in their data warehouses but leave empty any tables for which the CBO does not collect relevant data, so queries that reference those tables do not fail. The tables listed as high priority are required for CODI's primary functions, while tables listed as lower priority may be relevant but are optional.

- High Priority
  - CENSUS_LOCATION
  - DEMOGRAPHIC
  - ENROLLMENT
  - PREGNANCY
  - PREGNANCY_OUTCOME
  - PROGRAM
  - PROVIDER
  - PRO_CM
  - SDOH_EVIDENCE_INDICATOR
  - SESSION
- Low Priority
  - ASSET_DELIVERY
  - COST
  - CURRICULUM_COMPONENT
  - REFERRAL

3.5.3 Government Benefits Organizations

Government benefits organizations include any governmental organization that provides assets to individuals. In CODI@NC for 2022, the only government benefits organization is the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).

Government benefits organizations should create all RDM tables in their data warehouses and populate those tables to the best of their ability. If a table cannot be implemented, the data partner should create the table and leave it empty so queries that reference those tables do not fail.

- High Priority
  - CENSUS_LOCATION
  - DEMOGRAPHIC
  - DIAGNOSIS

[10] Additional data tables may become relevant to community-based organization partners after the initial pilot demonstration. For example, some community-based organization partners might capture VITAL signs, such as height and weight.

  - ENCOUNTER
  - ENROLLMENT
  - LAB_RESULT_CM
  - PREGNANCY
  - PREGNANCY_OUTCOME
  - PRESCRIBING
  - PROCEDURES
  - PROGRAM
  - PROVIDER
  - PRO_CM
  - SDOH_EVIDENCE_INDICATOR
  - SESSION
  - VITAL
- Low Priority
  - ALERT
  - ASSET_DELIVERY
  - CONDITION
  - COST
  - CURRICULUM_COMPONENT
  - FAMILY_HISTORY
  - IMMUNIZATION
  - REFERRAL
  - SESSION_ALERT

4. Specific Guidance

The CODI RDM represents clinical encounter information by incorporating parts of the PCORnet CDM, hereafter referenced as CDM. CODI augments the CDM table subset with ancillary tables to represent community-based health intervention programs and program participation. Together, the identified subset of CDM tables and the CODI ancillary tables comprise the CODI RDM. The RDM captures individual-level information that researchers, program evaluators, and other CODI users wish to explore to understand and improve interventions for community health.

The CODI RLDM supports identity management needed to link information on clinical encounter and program participation information from multiple data partners to the same individual. The RLDM incorporates privacy tables from the CDM and CODI's anonymous link tables to manage identity while preserving privacy.

Section 4.1 provides CODI-specific implementation guidance for CDM tables used by CODI. Section 4.2 provides implementation guidance for each of the CODI-specific ancillary tables that augment CDM to complete the CODI scope of functionality. Within both sections, data tables and attributes appear in ALL CAPS; see Appendix C and Appendix D for a complete dictionary of these tables and attributes.

4.1 PCORnet CDM Data Tables

This section provides CODI-specific implementation guidance needed for a few PCORnet CDM tables that CODI relies on. The CODI implementer should first consult the primary implementation guidance in PCORnet Common Data Model v.6.0[11] for all CDM tables that CODI uses, and then follow the supplemental guidance provided here for CODI specific refinements.

CODI relies on a subset of CDM tables, and in some cases, a subset of those tables' attributes. All of the CDM tables and attributes that CODI relies on are listed in the data dictionaries in Appendix C and Appendix D for the convenience of the reader, but not as the official source of record for CDM tables.

If the CODI network wishes to utilize future updates to CDM in CODI, then they must request an update to the CODI DM IG to reflect the CDM version change.

CODI supplemental guidance is provided for the following CDM tables listed in alphabetic order: CONDITION, DEMOGRAPHIC, ENCOUNTER, IMMUNIZATION, and PRO_CM.

4.1.1 CONDITION

The CONDITION table contains a single record for each condition, problem, or disease that an individual reports directly to a health professional or CBO. Whereas a record in DIAGNOSIS indicates the results of diagnostic processes and medical coding within healthcare delivery, CONDITION information is an informal reporting of a problem. CODI only requires two types of CONDITION data to support the CCWG's designated research scope.

[11] https://pcornet.org/wp-content/uploads/2020/12/PCORnet-Common-Data-Model-v60-2020_10_221.pdf

The first type of information needed is any non-clinical reporting of an individual having had COVID-19. The second is any reporting of health risks related to social determinants or social factors, for example an International Clinical Diagnosis (ICD) code of "Z59.1" meaning "Inadequate housing." Any report of ICD Z codes pertaining to social determinants of health as a condition is of importance to CODI.

All other condition data are not required for the CODI scope (anticipated queries) designated by CCWG for this version of the CODI data models. Other condition data may be provided if an organization chooses to do so.

4.1.2 DEMOGRAPHIC

The DEMOGRAPHIC table contains one record for an individual. This table should include information for children and adults with the age of two or above[12] and with at least one clinical visit or program participation record since the clinical-community distributed data network 's start date. Implementers should not include individuals without other records in the RDM. For example, an individual who has no encounter or program participation information, **and** no other clinical or self-reported health or social status information, should not be included in the DEMOGRAPHIC table. Individuals should be included in the DEMOGRAPHIC table if, for example, they have ENCOUNTER data, even if they do not have VITAL data, or the other way around—VITAL data without ENCOUNTER data (explained further in 4.1.3).

To preserve referential integrity, there must be a DEMOGRAPHIC record for any individual for whom information exists in any other RDM table (such as ENCOUNTER or SESSION). Conversely, every DEMOGRAPHIC record should have corresponding records in at least one other RDM table.

CODI omits individuals without other information in the RDM because adding individuals to the DEMOGRAPHIC table without sufficient information to answer possible research questions introduces an unwarranted risk. For example, an individual might be selected as a member of a cohort based on age and sex, but absent any encounters, vital signs, or program participation, none of the CODI research questions benefit from the inclusion of that individual. Although the CODI PPRL strategies are designed to mitigate the risks of sharing health information across organizations for research, those risks are not necessary for individuals who do not satisfy the research needs.

The SEX field is officially intended for sex at birth, however many of the community based data owners do not collect sex at birth, and so can populate this filed with the sex on record. GENDER_IDENTITY field is included in cases where the data owner collects both sex at birth and gender identity and wants to distinguish between the two. GENDER_IDENTITY is optional.

4.1.3 ENCOUNTER

The ENCOUNTER table contains a single record for each unique encounter. Several other CDM tables contain optional foreign key references to the ENCOUNTER table, including DIAGNOSIS, CONDITION, LAB_RESULT_CM, PRESCRIBING, PROCEDURES, and

[12] The growthcleanr package for cleaning longitudinal anthropometric observations is designed and optimized for evaluating height and weight measurements taken for subjects aged 2-65. However, data on individuals above 65 are allowed in CODI.

VITAL. These references are optional so that diagnoses, conditions, lab results, prescriptions, procedures, and vitals can be captured in the CDM even if there is no information available for any corresponding encounter or the corresponding encounter is unknown. In some cases, there may not even be a corresponding encounter, such as when vital signs are measured outside a clinical setting.

The CODI's ancillary table SESSION supplements the CDM ENCOUNTER table by representing information on encounters that involve screening or interventions with a focus on chronic disease prevention or management, including obesity, obesity prevention, healthy eating, or active living. This means that an ENCOUNTER record and a SESSION record may be linked. See the SESSION table implementation guidance for more detail.

4.1.4 IMMUNIZATION

The IMMUNIZATION table contains one record for each vaccination encounter or report for an individual regardless of whether the vaccination encounter is reported in ENCOUNTER or PROCEDURE. Vaccination status information may come from within the health system or elsewhere (including self-reported vaccinations and information from a vaccine registry).

To support the CCWG's designated scope of data queries for this version, only IMMUNIZATION records conveying vaccination status pertaining to the COVID-19 disease (e.g., the coronavirus SARS-CoV-2 vaccine) are needed. Other immunization data may be provided if an organization chooses to do so.

4.1.5 PRIVATE_DEMOGRAPHIC and PRIVATE_ADDRESS_HISTORY

The PRIVATE_DEMOGRAPHIC and PRIVATE_ADDRESS_HISTORY tables are used in the RLDM to temporarily contain information used in creating CODI record identification links. These tables are to be populated in a secure repository that the PPRL process can access but will not be accessible by other users of the RDM or by anyone who does not own the private demographic and address data.

These private tables contain the protected PII that the CODI implementer will obfuscate using a cryptographic hash function [13] that generates deidentified hash bundles of the PII for each individual. Only the obfuscated PII in the hash bundles is shared with a linkage agent as part of the record linkage process. The plain text PII data is separated and protected for the time it exists to create the hash bundles. The CODI Privacy Preserving Record Linkage Implementation Guide describes in detail the procedures performed by a CODI implementer to encrypt the data and the steps performed by the linkage agent to generate link IDs.

The CODI implementer will need to populate the private tables and then delete the data from the private tables once the LINKIDs are created. Refer to the PPRL IG to learn more about the in-between steps.

A data owner will provide the private information to the CODI implementer using a sharing mechanism that both parties agree upon and that keep the PII secure until it is deleted in the final PPRL process step (if the data owner is a CODI implementer, this is an internal process).

[13] https://en.wikipedia.org/wiki/Cryptographic_hash_function

CODI's PPRL process requires just one demographic record and one address record per individual per data source. This means that the implementer who is populating the CODI RLDM private demographic table must ensure that individuals (patients or participants) are not duplicated in the CODI RLDM private demographic table.

The SEX field is officially intended for sex at birth, however many of the community based data owners do not collect sex at birth, and so the implementer shall populate this field with the sex on record.

The PPRL tool requires only one address per individual. However, the RLDM private address history table may contain multiple addresses per individual used for populating an individual's Census location history in the RDM. In the case of multiple addresses per individual, each record must have an ADDRESS_PERIOD_START value and except for the record of the individual's current address, each record must have an ADDRESS_PERIOD_END value. For the case where the start date is not available, see Appendix B for the CODI network-specific decision on a default value for ADDRESS_PERIOD_START.

If more than one address is present in PRIVATE_ADDRESS_HISTORY, the PPRL tool will select the record with a 'Y' value in ADDRESS_PREFERRED. Ideally, this address will be an individual's most current, valid address. When multiple addresses are present for an individual, the implementer shall ensure that only one address is marked as preferred.

For detailed instructions on which address is preferred for PPRL, follow these rules as closely as possible:

- Using the terminology of the PCORnet private address history, select the record with the latest address period start-date and null address period end-date, or if no null end-date, then select the address with the latest period end-date. Select an address where the address usage is the home address. If the address preferred flag is set to a non-null value, select a record where the address preferred flag is true. Finally, select a record where the address type is physical or both physical and postal.
- If the data owner does not already use a PCORnet PRIVATE_ADDRESS_HISTORY table, then use the address from the owner's native representation selecting the address that approximates the same rules as described above. Transform the native address as needed to map to the RLDM PRIVATE_ADDRESS_HISTORY table.

The private table fields used by PPRL are listed below. While the private tables in the PCORnet CDM contain other fields, CODI does not use them in the PPRL process.

Table 4: Fields used by the PPRL process

Private Address History

CODI Attribute	Cardinality	Type
PATID	1	FK::DEMOGRAPHIC
ADDRESS_STREET	0..1	String
ADDRESS_DETAIL	0..1	String
ADDRESS_ZIP5	0..1	CHAR (5)
ADDRESS_PREFERRED	1	CHAR (1) {Y or N}

Private Demographic

CODI Attribute	Cardinality	Type
PATID	1	FK::DEMOGRAPHIC
PAT_FIRSTNAME	0..1	String
PAT_LASTNAME	0..1	String
BIRTH_DATE	0..1	date
SEX	0..1	SEX_TYPE
PRIMARY_PHONE	0..1	CHAR (10)

Prior to executing the PPRL process, implementers should ensure that the PRIVATE_ADDRESS_HISTORY records are the highest quality possible. Implementers can achieve this by applying a geocoding function on address values. The function will standardize the address data and flag addresses that may not exist.

The addresses are also used for identifying an individual's Census location in the CODI RDM. The implementer shall convert addresses to Census geocodes to provide general residential areas for all individuals. The CODI RDM table, CENSUS_LOCATION can capture an individual's location geocode for a current address as well as past addresses. This will support queries based on area-level Census data (See Section 4.2.3).

After PPRL steps are executed, and the link ids, household ids, and geocodes are established, delete the contents of the PRIVATE_ADDRESS_HISTORY and PRIVATE_DEMOGRAPHIC tables to minimize unintentional disclosures.

4.1.6 PRO_CM

The CDM Patient-Reported Outcome Common Model (PRO_CM) table is used to store responses to patient-reported outcome measures (PROs) or questionnaires. This table can be used to store item-level (i.e., single question) responses as well as the overall score for each measure associated with a related set of questions, for example in a screening segment.

To support the CCWG's designated scope of data queries for this version, only data for questions or measures relevant to understanding an individual's social circumstances are requested. These social determinants of health (SDOH) include access to adequate food, housing, transportation, and personal safety. Other social circumstances also apply. Data owners are encouraged to share any screening responses or measures for any SDOH using the PRO_CM table. The implementation instructions are in the PCORnet CDM specification.

CDM's guidance for PRO_CM permits several alternative mappings from the data owner's source data. Refer to CDM IG for examples. CODI implementers should follow CODI's more specific guidance provided in the CODI DM IG, here, to maintain consistent use of PRO_CM across data owners in a CODI network.

The CODI specification interprets 'items' in PRO_CM as question-response pairs. These may appear in a questionnaire form, survey instrument, or in a clinical flow sheet and may be administered verbally, or performed by the individual electronically, or on paper. CODI accepts any question-response pairs in the PRO_CM table. The question is mapped to a PRO_ITEM. The response is mapped to PRO_RESPONSE field.

CODI interprets a 'measure' in PRO_CM as a set of patient-reported evidence which is evaluated by a health practitioner or is calculated from the responses of a group of items (e.g., questions). A measure might correspond to an entire questionnaire containing closely related questions, or to a section of that questionnaire which when evaluated produces a score.

Table 5 provides specific CODI guidance on PRO_ITEM and PRO_MEASURE fields. For the comprehensive table dictionary, see the DM IG appendix, For PCORnet guidance see the CDM IG.

Table 5: CODI Guidance for PRO_ITEM and PRO_MEASURE Fields

Column	CODI Guidance
PRO_TYPE	'LC' if the question is LOINC coded, or 'OT' (i.e., Other) if the question is not coded
PRO_ITEM_LOINC	Question's LOINC code if applicable, otherwise NULL
PRO_ITEM_NAME	A cryptic unique name for the question. For example, an EHR system flowsheet number such as '1570008807'. This is optional if the question is coded. If it is not coded, it is recommended to have a short name for reports.
PRO_ITEM_FULLNAME	A more understandable name for the question. For example, 'CHILDREN'S HEALTHWATCH HOUSING HOMELESS'. This is an optional field.
PRO_ITEM_TEXT	The exact text of the question. For example, 'In the last 12 months, was there a time when you did not have a steady place to sleep or slept in a shelter (including now)?'
PRO_RESPONSE_TEXT	The participant's textual or coded response to the question. If the response is coded, then populate this field with that code. The PRO_TYPE field above will name the code system and the end user can look up the associated textual answer (e.g., 'Rarely') in the code system's look-up table. If the answer is not coded, and it is a textual answer, then populate the field with the response text.
PRO_RESPONSE_NUM	Only use this field for numeric responses as with a question that asks, 'How many alcoholic beverages do you consume per week', where the answer is a number, e.g., 5.
PRO_MEASURE_NAME	If the PRO_ITEM referred to in PRO_ITEM_TEXT, is a question in a form or question group, you may populate this field with the name of the form or group, if the source system has a name.
PRO_MEASURE_FULLNAME	If the PRO_ITEM referred to in PRO_ITEM_TEXT is a question in a form or question group, has a lengthy name in addition to a cryptic name (which goes in the PRO_MEASURE_NAME), then populate this field with the lengthy name of the form or group.
PRO_MEASURE_SEQ	If the PRO_ITEM referred to in PRO_ITEM_TEXT, is a question in a named or coded form or question group, you may populate this field with the sequential order of the PRO_ITEM as it appears in the group or form.
PRO_MEASURE_LOINC	If the PRO_ITEM referred to in PRO_ITEM_TEXT, is a question in a LOINC coded form or question group, you may populate this field the LOINC code for that form or group.
PRO_MEASURE_VERSION	If the PRO_ITEM referred to in PRO_ITEM_TEXT, is a question in a coded or named form or question group, and it has an associated version number, date, or other string, then you may use this field.

4.2 CODI Ancillary Tables

The ancillary tables are designed to augment the CDM so that CODI research inquiries that are not supported in CDM are supported in the RDM. There are 14 ancillary tables designed specifically for CODI and three tables adopted from other models but ancillary to CDM.

ALERT
ASSET_DELIVERY
CENSUS_LOCATION (from CHORDS)
COST (from OMOP)
CURRICULUM_COMPONENT
FAMILY_HISTORY
HOUSEHOLD_LINK
LINK
PREGNANCY
PREGNANCY_OUTCOME
PROGRAM
PROGRAM_ENROLLMENT
REFERRAL
SDOH_EVIDENCE_INDICATOR
SESSION
SESSION_ALERT

4.2.1 ALERT

The ALERT table contains one record for each distinct **kind** of alert directly related to chronic disease, especially cardiometabolic-related diseases. Each CODI implementer will determine on behalf of their organization which alerts qualify. For each such alert, the ALERT table captures information about the circumstances surrounding that alert. ALERT is a reference table that will likely need to be populated manually because the information it contains requires human curation. The attributes appearing in this table are intended to help a researcher understand when and why an alert might trigger.

Once CDS Hooks[14] (or a similar standard) becomes more widely adopted, the ALERT table should be updated to reflect that standard rather than relying solely on prose documentation and human identification of relevant alert types.

4.2.2 ASSET_DELIVERY

The ASSET_DELIVERY table contains one record for each contiguous time period during which a person consistently receives assets. An asset is a resource transferred by a program to an individual. The intention is that each record represents a series of asset deliveries that regularly transpires. In situations where each delivery is ad hoc, the expectation is that a separate record appears for each such delivery. Otherwise, CODI assumes the deliveries occur on a recurring basis as described by the record. DELIVERY_FREQ indicates the number of deliveries within each unit of time. DELIVERY_FREQ_UNIT establishes the corresponding unit of time.

Monthly refers to calendar months. Deliveries that happen every 28 days should be encoded as 0.25 deliveries every week (i.e., once every four weeks). For example, an individual might receive cash benefits twice every calendar month. The start and end dates indicate the period

[14] https://cds-hooks.org/

during which these benefits were received, with a DELIVERY_FREQ_UNIT of monthly and a DELIVERY_FREQ of 2.

CODI implementers acting on behalf of data owners that participate in asset delivery are encouraged to populate the ASSET_PURPOSE at a minimum because it provides researchers with insight into the circumstances surrounding the delivery of assets.

4.2.3 CENSUS_LOCATION and Census Data

The CENSUS_LOCATION table incorporated from the CHORDS VDW data model, links an individual to a geographic area that corresponds to the individual's residence. Every individual in DEMOGRAPHICS must have at least one record in CENSUS_LOCATION with a Census location code (GEOCODE)[15] field value, and a start date (LOC_START) field value that corresponds to the earliest known date for which the location code is valid.

Ideally, the LOC_START value is determined from the earliest encounter between the data owner and the individual in which an address was captured. See Appendix B for any CODI network-specific decisions about populating the LOC_START field when the location start date is not known.

CODI implementers shall assign a geographic area to an individual based on the individual's geocoded[16] private home address. The Census Bureau creates and maintains geographic codes for statistical geographic areas including census tracts. The geocode combines both the Federal Information Processing Standards (FIPS) codes and Census Bureau codes to represent nested levels of geographic areas (state, county, census tract, …). Implementers use an individual's private address data as input to a geocoding application executed in a secure system separate from shared CODI data.

The preferred geographic area is a census tract, however, if that geographic specificity cannot be determined, then provide a county or state level geographic area code, using FIPS. Census block group and block level areas are too geographically specific risking privacy and shall not be assigned to individuals in CENSUS_LOCATION. If the individual's location is unknown, then insert a record for that individual with a LOC_START value and an empty GEOCODE field.

If a data owner has past residence(s) data on an individual, those locations can be added to provide CODI users a location history. The 'current' known location for an individual is identified by a missing location end date (LOC_END), while a populated end date field represents a past location. The historical records should have start and end dates that convey the temporal order of an individual's location.

The data coordinating center (DCC) is responsible for maintaining a reference table containing geographic area (census tract) level demographic data that can be used for distributed project queries. Those data may be used as cohort filters or covariates as defined by a specific project and data use agreement. Data owners are not expected to maintain population-level demographic data, but if it is determined that such data would be useful for local/internal queries, they may

[15] https://www.census.gov/programs-surveys/geography/guidance/geo-identifiers.html

[16] A string up to 15 digits long (11 for the tract level) with geographic information hierarchically documented through the string of digits. A numeric string up to 15 digits, of multiple census variables for geographic information hierarchically documented through the string of digits: state(2) + county(3) + tract(6), blockgp(1) + block(3) = GeoCode(15)

request the most current population-level demographic reference table from the DCC which can be stored within the CODI data mart. Area level demographic data will be drawn from sources such as American Community Survey (ACS)[17], which is constructed by the Census Bureau and made publicly available. The DCC will also manage any population-level statistics associated with geo-locations that are needed in CODI queries.

4.2.4 COST

The OMOP COST table captures records containing the cost of any medical event or program participation occurrence recorded in CODI RDM tables. This table does not capture the cost of providing the service, but rather the amounts billed and received. The COST table can link to ENCOUNTER, LAB_RESULT_CM, PROCEDURES, or SESSION.

4.2.5 CURRICULUM_COMPONENT

The CURRICULUM_COMPONENT table enumerates the standard elements of a program. It supports both a fixed curriculum, in which the components are ordered using SESSION_INDEX, and a recurring curriculum, in which the components repeat. Repeating components are documented with a combination of SESSION_FREQ and SESSION_FREQ_UNIT, as described above.

The remaining attributes mirror those in the SESSION table (as described below). The CURRICULUM_COMPONENT table describes what is intended to happen throughout the course of the program. The SESSION table describes what has been documented as having transpired. The CURRICULUM_COMPONENT table provides researchers with insight into what likely happened when session information is missing or incomplete.

4.2.6 FAMILY_HISTORY

The FAMILY_HISTORY table stores information regarding an individual's family health. Each entry records a single condition reported for a family member. Thus, if an individual's parents both have a history of chronic disease, two records would be present in this table. The intention is that CODI implementers only retrieve family history information present in a patient's or program participant's own record. If the EHR provides links to a parent's medical record, that information should **not** be included in FAMILY_HISTORY. Reported conditions must be linked to controlled vocabulary—an ICD-9, ICD-10, or Systematized Nomenclature of Human Medicine (SNOMED) code—so researchers can easily interpret the reported family condition. Implementers will need to map from whatever terminology is used for family history to one of these vocabularies.

4.2.7 HOUSEHOLD_LINK

The HOUSEHOLD_LINK table contains one record for each person in the demographics table for each iteration of the record linkage process. When the RLDM is initially populated, this table will be empty. Each time the record linkage process is completed is a distinct iteration. For each

[17] ACS is an ongoing survey, conducted annually by the United States Census Bureau, which provides vital information about the United States population. The ACS has an annual sample size of about 3.5 million addresses, with surveys administered and information collected nearly every day of the year. Data are pooled across a calendar year to produce annual estimates. https://www.census.gov/programs-surveys/acs/

iteration, the linkage agent will provide each CODI implementer with the information needed to establish an iteration-specific mapping from HOUSEHOLDID to PATID. Upon receipt of this mapping, implementers should populate the HOUSEHOLD_LINK table with contents from the mapping for that iteration. The implementer should leave previous iterations in the HOUSEHOLD_LINK table for managing past queries and results from previous iterations (assuming an individual's PATID remains the same from iteration to iteration).

Each household link record includes a designated household. More than one record may have the same household designation, indicating multiple people belong to the same household. The household designation indicates the existence of a household address where one or more persons in the demographics table are believed to reside for a given iteration of household linkage.

The linkage agent aligns households across partners and generates corresponding household link IDs using obfuscated information from the PRIVATE_ADDRESS_HISTORY table. The plain text household address is not shared with the linkage agent, nor is it given a CENSUS_LOCATION. Only an anonymized household identifier in the HOUSEHOLD_LINK record conveys the household to which a person is linked.

A household does not have an enduring ID; it is changed for each iteration of the record linkage process. There is no longitudinal information on a household; a user cannot track changes in household composition. Finally, a household says nothing about relationships (e.g., child, spouse, roommate) among household members, only that the household member is co-located with other members based on the address information available at the time of the record linkage iteration.

4.2.8 LINK

The LINK table contains one record for each person in the DEMOGRAPHIC table for each iteration of record linkage. When the RLDM is initially populated, this table will be empty. The LINKID is populated as part of the record linkage process. Each time the record linkage process is completed is a distinct iteration. For each iteration, the linkage agent will provide the information needed for the CODI implementer to establish an iteration-specific mapping from PATID to LINKID for each data owner. Upon receipt of this mapping, the implementer should populate the LINK table with contents from the mapping for that iteration. The implementer should leave previous iterations in the LINK table for managing past queries and results from previous iterations (assuming an individual's PATID remains the same from iteration to iteration).

4.2.9 PREGNANCY

The PREGNANCY table contains one record for each time an individual is pregnant. It stores information about the pregnant person and the circumstances of the pregnancy. For example, it provides a single place for information about the pregnant person's weight, body mass index (BMI), use of tobacco, use of alcohol, and prenatal care. This table was designed to be compatible with WIC information systems.

Much of the information about an individual's pregnancy is captured in CDM. The PREGNANCY table consolidates that information into a single location. Implementers should only populate this table if they can do so with certainty—clinical implementers are likely to find that pregnancy information cannot feasibly be assembled from the EHR. For example,

CIGARETTE_FIRST captures information about tobacco usage during the first trimester. If an implementer is not certain that its information about tobacco usage is specific to the first trimester, that implementer should leave that attribute blank.

4.2.10 PREGNANCY_OUTCOME

The PREGNANCY_OUTCOME table contains one record for each infant resulting from a given pregnancy. It stores information about the individual(s) resulting from the pregnancy. It consolidates information about the child's height and weight at birth, breastfeeding, and exposure to tobacco. This table was designed to be compatible with WIC information systems. As with PREGNANCY, clinical implementers are likely to find that this table cannot feasibly be populated.

If an implementer has multiple values for an attribute in this table, the most current datum should be selected. For example, if the parent has reported about breastfeeding multiple times, the most recent breastfeeding information should appear in PREGNANCY_OUTCOME.

To maintain referential integrity, each record in the PREGNANCY_OUTCOME table must link back to a corresponding record in the PREGNANCY table. A link back to the DEMOGRAPHICS table should also appear, **if** the implementer has a record for the child. If the normal demographic information is not available for the child, then the PATID fields should be left blank.

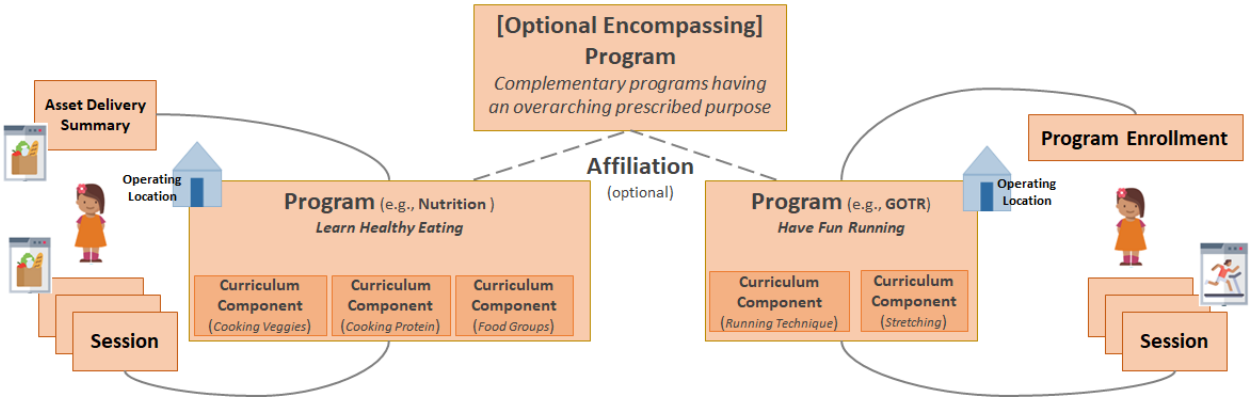
4.2.11 PROGRAM

The PROGRAM table contains one record for each distinct chronic disease-related program. For the purposes of CODI, each location at which a program is administered constitutes a distinct program. For example, each clinic that administers a weight management program appears separately in the PROGRAM table.

This is the second table that will likely need to be manually populated. It captures a program manager's best understanding of how a program, for example weight related or nutrition related is administered and for what purpose. The attributes with the PROGRAM_ and AIM_ prefixes apply to every program. The attributes with the PRESCRIBED_ prefix only apply to those programs with a predefined frequency of interaction, such as a program that lasts for 10 weeks and meets twice a week, two hours each time. This regularity allows researchers to know the intended dose and intensity (i.e., frequency of interaction) for the program. Programs without a predefined dose should leave these attributes blank.

The AFFILIATED_PROGRAM attribute provides a way to document that a given program is affiliated with an encompassing program or program category. For example, consider a weight-related program with two component programs (a cooking class and a physical activity program); participation in each is based on each individual's needs: this configuration includes three programs. The affiliated programs (i.e., cooking class, physical activity program) include prescribed doses and have specific aims, while the parent program has no set dose, and its aims are broad. The AFFILIATED_PROGRAM attribute allows the affiliated programs to indicate the encompassing program, or program category with which they are affiliated.

Figure 4. Illustration of Affiliated Programs



The attributes with the LOCATION_ prefix describe the location at which the program is typically administered. Three variants of location are supported: address, geospatial coordinates (latitude and longitude), and geocode (typically census tract). Implementers should provide all three variants for each program they are extracting, if possible.

Some programs are designed to be attended online or at home with program materials. The LOCATION_ADDRES for these programs should be indicated with the string ‘Virtual’.

The attributes with the SESSION_OMISSION_ prefix describe the circumstances under which session information is missing for the program. For example, some clinical programs only record sessions with a clinical component. The sessions lacking a clinical component are not documented and therefore do not appear in the SESSION table. Other programs exhibit less systematic omissions (e.g., because attendance is sometimes captured on paper). These attributes are included to help researchers better decide how to handle missing session information.

4.2.12 PROGRAM_ENROLLMENT

The PROGRAM_ENROLLMENT table captures information about an individual’s enrollment in a chronic disease-related program or specific program provided by a CBO (see PROGRAM, above). Enrolling in a program does not mean the individual participated in the program, only that they were registered to participate. The individual’s participation is captured in the SESSION table if attendance or encounter information is available. The PROGRAM_ENROLLMENT table also captures the individual’s completion of the program. How a program defines completion is left to the discretion of each program. Researchers should consult the program’s documentation to understand how best to interpret program completion. If an individual disenrolls from the program, implementers should populate the disposition description to document the circumstances involved.

4.2.13 REFERRAL

The REFERRAL table contains one record for each outgoing or incoming referral for clinical or program services. The DIRECTION attribute indicates if the record represents a data owner organization initiating a referral (outgoing) or receiving a referral (incoming). Internal referrals should result in two records in the REFERRAL table: one outgoing referral and a second

incoming referral. The purpose of the source and destination organization attributes is to link outgoing referrals with incoming referrals so researchers can see whether a referral successfully connected an individual with a weight-related or other health-related program. Implementers will need to map source and destination organizations to CMS Certification Numbers where possible; see ORGANIZATION_TYPE in Appendix D for more information about coding organizations.

4.2.14 SDOH_EVIDENCE_INDICATOR

Social determinants of health (SDOH) are conditions that can affect a wide range of health risks and outcomes. Example social determinants include situations concerning housing, food, and personal safety. The ability to analyze SDOH data is key to understanding and achieving health equity.

A record in the CODI SDOH_EVIDENCE_INDICATOR table signals that data pertaining to the personal-level social conditions of an individual are available through CODI. SDOH screenings (e.g., surveys, questionaries) data, and self-reported conditions or professionally assessed clinical diagnoses (e.g., using ICD, SNOMED, or some other coding system) are examples of individual-level SDOH information. Another example of SDOH evidence is an individual's participation in a program whose enrollment is due to certain social circumstances, for example, homelessness.

Domains in which SDOH evidence can be categorized are listed below and defined in the CODI SDOH_CATEGORY_TYPE. See Table 81 for category definitions. For an explanation of the origin of these domains, see Appendix A, CODI SDOH Categories and The Gravity Project.

High Priority

- FINANCIAL_DOMAIN
- FOOD_DOMAIN
- HEALTH_INSURANCE_DOMAIN
- HOUSING_ADEQUACY_DOMAIN
- HOUSING_STABILITY_DOMAIN
- INTERPERSONAL_VIOLENCE_DOMAIN
- TRANSPORTATION_DOMAIN

Low Priority

- EDUCATION_DOMAIN
- ELDER_CARE_DOMAIN
- EMPLOYMENT_DOMAIN
- MATERIAL_NECESSITIES_DOMAIN
- SOCIAL_CONNECTION_DOMAIN
- STRESS_DOMAIN
- VETERAN_DOMAIN

It is important to understand that an individual with an associated SDOH evidence indicator, does *not* imply whether the indicated evidence posits a negative (insecure) or positive (secure) social condition; it merely signals to the CODI user that some CODI data exists that can provide some social condition information for that individual. For example, consider the screening question,

> Within the past 12 months, have you been unable to get utilities (heat, electricity) when it was really needed?

An individual's response to this question, no matter the response value (e.g., 'Yes', 'No') is evidence on housing adequacy for this individual. Therefore, the evidence is captured in screening data, flagged with a record in the SDOH_EVIDENCE_INDICATOR table, and linked to that individual.

Understanding the magnitude of this individual's housing security problem (or lack thereof) requires the data user to examine all the data that all the housing evidence indicators point to, and then consider that evidence in the context of the research question.

Purpose of the SDOH_EVIDENCE_INDICATOR

Records in the SDOH_EVIDENCE_INDICATOR table provide CODI users a shortcut to CODI data containing any SDOH evidence on an individual. The record includes the SDOH category of the evidence, and the CODI table and record in which the evidence can be found. This shortcut lets users filter individuals based on the presence and category of evidence.

The SDOH_EVIDENCE_INDICATOR postpones the need for the CODI user to address the complex and disparate ways in which each data owner represents individual-level social determinants data allowing the data user to scan across data from multiple data owners and different SDOH collection practices for an initial selection of individuals to study. CODI captures specific detailed evidence in its various forms in CODI tables such as PRO_CM,[18] DIAGNOSIS, and CONDITION and captures each data owner's choice of codes, text, and values.

CODI's approach allows clinical and community organizations to contribute whatever SDOH information their organization has, with minimum modification, and still give users a harmonized, preliminary view of available SDOH data.

Implementation Guidance for SDOH_EVIDENCE_INDICATOR

The SDOH_EVIDENCE_INDICATOR table contains zero or more records for each PATID in the DEMOGRAPHICS table. A record must contain a PATID, an SDOH_CATEGORY field value from SDOH_CATEGORY_TYPE, and either an EVIDENCE_TABLE value matching a CODI table name (if applicable), or an EVIDENCE_EXPLANATION value (or both).

An individual may have more than one record with the same SDOH category because there may be more than one piece of SDOH evidence in that category for that individual (captured in different tables and rows).

CODI implementers shall insert SDOH evidence indicator records based on evidence in the data owners' extracted dataset. While most SDOH data expected from CODI data owners map to a CODI table, there are some cases where the SDOH evidence does not explicitly map to a CODI table. In these cases, the data owner's evidence should be extracted into the dataset and transformed into a textual explanation for the EVIDENCE_EXPLANATION field, as a substitute for referencing evidence in any CODI table.

There are several source scenarios for SDOH evidence, each requiring a slightly different data ETL process. Table 6 lists the possible scenarios and their associated CODI tables (if applicable). Subsequent sub-sections provide implementation guidance for each scenario in this table.

[18] Patient Reported Outcome Common Model (PRO_CM) enables the storage of any question or data element and any answer or value, coded or not, and standard or not. See Section 4.1.6 for more explanation.

Table 6: SDOH Data Source Scenarios

#	Possible SDOH Data Scenario	Associated CODI Table(s)
1	Screening questions and answers (LOINC[19] coded or not)	PRO_CM
2	Diagnosis or condition codes (e.g., ICD[20] Z59.01 Sheltered homelessness, SNOMED-CT[21], or other coding systems)	DIAGNOSIS, CONDITION
3	Enrollment and participation in qualifying service programs (e.g., food assistance programs)	PROGRAM_ENROLLMENT, SESSION
4	Material or monetary assistance through various community-based programs	ASSET_DELIVERY
5	Program enrollment fee waivers and other information	EVIDENCE_EXPLANATION field in SDOH_EVIDENCE_INDICATOR.

1. SDOH screening questions and answers

For each social determinants screening item added to the PRO_CM table, the implementer shall insert one or more records into the SDOH_EVIDENCE_INDICATOR table using the applicable SDOH_CATEGORY_TYPE name. Populate the field, EVIDENCE_TABLE with 'PRO_CM' and EVIDENCE_ROWID with the PRO_CM_ID (primary key) for the row in that table. The EVIDENCE_EXPLANATION is not necessary for this source scenario.

To automate this process, implementers should pre-map the data owner's standard SDOH screening instruments, by question, to the appropriate SDOH category for reference during ETL processing. For example, the implementer should map all the data owner's housing instability and homelessness-related questions to the HOUSING_INSTABILITY_DOMAIN category, and likewise with all screening questions, pairing each screening question with at least one category (if appropriate). Some questions may fall into more than one category and should be mapped to each of the relevant categories.

The Gravity Project, a consortium working to harmonize social risk factor data, has already categorized many standard screening tools and their questions by SDOH domains. Refer to the Gravity Terminology Workstream[22] to view the list of mapped questions and codes for each domain. For an explanation of the Gravity Project and CODI SDOH domains, see Appendix A.

The data owner and implementer shall decide on an appropriate CODI SDOH category for any SDOH screening question that is not already categorized by the Gravity Project, to include in the question-to-category mapping reference.

Using the question-to-category mapping reference, the CODI implementer will design the ETL process to enter one or more SDOH_EVIDENCE_INDICATOR records with the dictated SDOH

[19] Logical Observation Identifiers Names and Codes (LOINC)
[20] International Classification of Diseases (ICD)
[21] Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT)
[22] https://confluence.hl7.org/display/GRAV/Terminology+Workstream+Dashboard

category, for each question entered in the PRO_CM table, no matter the individual's response value.

Note that an individual can have multiple evidence indicator records in the same SDOH category, and for the same or different evidence tables. Each record must have a distinct combination of PATID, EVIDENCE_TABLE, EVIDENCE_ROWID, and EVIDENCE_EXPLANTION values.

2. Diagnosis or self-reported conditions

Clinical organizations may assign a social determinant diagnosis or condition code (e.g., an ICD Z code) to an individual which is then contained in the DIAGNOSIS or CONDITION table.

Implementers should create a diagnosis and condition code-to-SDOH category mapping reference for any SDOH diagnosis or condition codes that are used in the data owner's information system. For each entry in the DIAGNOSIS or CONDITION table that matches a code in the code-to-SDOH category mapping, the implementer should design the ETL process to also enter an SDOH_EVIDENCE_INDICATOR record with the appropriate field values.

3. Program enrollment and participation

Another type of SDOH evidence, which is less explicit than a diagnosis, condition, or screening question, is a social circumstance that is implied by an individual's participation in a social program for those in need. For example, an individual qualifying for and enrolled in a homelessness program implies some level of housing insecurity for that individual. In this scenario, an evidence indicator record shall be added for that individual in which the EVIDENCE_TABLE is PROGRAM_ENROLLMENT or SESSION and the EVIDENCE_ROWID is the unique ID to that individual's enrollment record.

Implementers should create a program-to-category mapping reference for use in the ETL process.

4. Material or monetary assistance

Some social programs provide assets (e.g., food vouchers) to their program participants. This data is represented in CODI's ASSET_DELIVERY table. Providing assets such as a food voucher to qualifying program participants may suggest a social circumstance that warrants an SDOH_EVIDENCE_INDICATOR. If so, the implementer shall add an indicator record with the appropriate SDOH category (e.g., FOOD_DOMAIN) for any individual with an ASSET_DELIVERY record having a qualifying ASSET_PURPOSE value.

Implementers should create an asset-purpose-to-category mapping reference for use in the ETL process. (See Table 45. ASSET_TYPE Details.)

5. Program fee waiver and other information

There are scenarios in which certain social circumstances are implied by relevant information in the data owner's information system, but not transferred explicitly to a CODI table. One example is whether a program participant is receiving financial aid or a waiver for participation fees, such as a Department of Parks and Recreation membership fee. CODI does not explicitly capture a fee

waiver (or any individual-level income) information. However, if this fee waiver is due to an individual's financial security, then this is SDOH evidence.

Data owners and implementers can decide to enter an evidence indicator for known circumstances captured in the owner's source system, but not captured in CODI tables, by using the EVIDENCE_EXPLANATION field of the indicator record. In this scenario, the implementer inserts an indicator record for PATID, with or without an EVIDENCE_TABLE value, with the appropriate SDOH_CATEGORY_TYPE (e.g., FINANCIAL_DOMAN), and uses the EVIDENCE_EXPLANATION field to explain the reason for the indicator.

CODI does not currently provide a standard set of string values for the evidence explanation data element, but the CODI implementation network membership can establish standard explanation values for their network.

This scenario works for a program fee waiver and discount if due to financial need but can also work for other information important to SDOH, but not explicitly captured in CODI tables. For this type of scenario, PAT_ID, EVIDENCE_EXPLANATION, and SDOH_CATEGORY_TYPE are mandatory fields. EVIDENCE_TABLE and EVIDENCE_ROWID are optional if the evidence is not explicitly in a CODI table.

4.2.15 SESSION

The SESSION table contains one record for each interaction between an individual and a healthcare provider or program representative. In its most basic incarnation, the SESSION table is an extension of the ENCOUNTER table, to include CODI-specific attributes. For example, during a well-child visit or adult's annual check-up, exercise and nutrition screening may transpire. When an ENCOUNTER involves multiple providers interacting with an individual, multiple SESSION records should be created. For example, a single encounter sometimes includes an individual interacting with multiple providers, such as the primary care physician and a dietician. Each of these interactions is a separate session because they involve different providers.

As an example of the difference between PROGRAM data and SESSION data, consider Figure 5. For Girls on the Run, there would be one entry for each school where Girls on the Run is offered. In the figure, only a single program record is shown (and many details, including location, are omitted for brevity). In the SESSION table, there is one record for each time an individual attends the program. In this example, there are two children (G234 and G567). The first child attended Girls on the Run three times, and the second child attended twice. The ENCOUNTERID is missing because Girls on the Run is a community program that does not collect clinical information.

Figure 5. Sample Program and Session Data

PROGRAM ID	PROGRAM_ NAME	PROGRAM_ SETTING	PROGRAM_ MODE	AIM_ NUTRITION	AIM_ ACTIVITY	AIM_ WEIGHT	TOTAL_ DOSE
001	Girls on the Run	CO	G	False	True	False	20.0

SESSION ID	PAT ID	ENCOUNTER ID	PROVIDER ID	PROGRAM ID	SESSION_ DATE	DOSE
001	G234		003	001	14-Jan-2019	1.0
002	G234		003	001	16-Jan-2019	1.0
003	G234		003	001	21-Jan-2019	1.0
004	G567		003	001	14-Jan-2019	1.0
005	G567		003	001	21-Jan-2019	1.0

At a minimum, implementers should populate the SESSION table for wellness visits and for encounters that are part of a chronic disease intervention program (e.g., Diabetes Prevention Program, MEND, Healthy Weight Clinic, Girls on the Run, or Hunger Free Colorado). The next highest priority is to populate the SESSION table for primary care encounters and for encounters related to chronic diseases or chronic-related comorbidities (e.g., a follow-up weight check, a visit for nutritional counseling, or a visit with a specialist such as endocrinology or cardiology). Finally, if the implementation budget allows, implementers should populate the SESSION table for every encounter type having session data

If the data owner or implementer chooses only some encounter types to extend with session data, then there should be no session records created for encounter types that are not one of those chosen encounter types. For example, if an implementer populates the SESSION table for only well visits and chronic disease intervention program encounters, then no record in the SESSION table should be created for emergency department encounters.

For programs related to chronic disease and based in a community setting, each SESSION record corresponds to an individual's participation in the program. For example, an individual who completes a program that meets weekly for 10 weeks should have 10 distinct SESSION records.

The DOSE attribute indicates the amount of time spent interacting with the individual (in hours). This attribute should only be populated based on what is documented in the EHR or other IT system. If the duration of the session is not documented, the DOSE attribute should remain empty. For example, the DOSE attribute would not be populated for interventions conducted by mail and may not be populated for Web-based interventions.

The SESSION table includes several process-related attributes (SCREENING, COUNSELING, and those with the INTERVENTION_ prefixes). In some cases, the values of these attributes need to be established based on local program knowledge as opposed to what is present in the

EHR. For example, if a program stipulates that every session includes physical activity, that attribute can be set solely based on attendance information, because the EHR or IT system may not track whether physical activity happened—it always happens.

4.2.16 SESSION_ALERT

The SESSION_ALERT table contains one record for each alert triggered during a session. In other words, it indicates an alert triggered in the context of a session. The intention is that a record in this table indicates that the provider responsible for a SESSION was made aware of a given ALERT.

5. Additional Resources

5.1 Requesting Changes

CODI Implementers and researchers are likely to find limitations with the Implementation Guide or CODI Data Models and can request changes. The following process will be followed to process those change requests.

The data owner, implementer, or researcher shall document the change request and send that request to CODI@cdc.gov.

MITRE determines which documents or data models, if any, might need to be changed.

MITRE presents the potential changes to the Scoping and Use Case Subgroup, also known as the Research Question Subgroup (within two weeks of the initial request) for feedback.

CDC decides how to handle the request based on the subgroup feedback.

MITRE implements any necessary changes and uploads the new documents to the MITRE external Microsoft Team's web site for the Implementation Subgroup.

CODI Teams site notifies subscribers of the availability of the updated documents.

Changes will follow typical semantic versioning. Changes that are backwards compatible increment the minor version number of the Implementation Guide (e.g., from 1.4 to 1.5). Changes that are not backwards compatible (e.g., replacing an ancillary table with a table from CDM) will increment the major version number (e.g., from 3.3 to 4.0). Changes in the implementation guidance that do not require any data model change such as revising guidance to improve clarity, will increment the iteration number of the minor version (e.g., from 4.1.2 to 4.1.3).[23]

5.2 Questions

Any questions regarding this implementation guide should be sent to CODI@cdc.gov.

[23] Because CODI Version 4.1 had not yet been deployed, there are some changes from 4.1 to 4.2 that are not backward compatible. For example, the table "ENROLLMENT" was renamed to "PROGRAM_ENROLLMENT". However, most changes from 4.1 to 4.2 were the inclusion of additional PCORnet tables and fields, and clarifications or corrections to the implementation guidance.

Appendix A CODI SDOH Categories and The Gravity Project

There are several local and national efforts to standardize the way clinical organizations and CBOs assess SDOH among the populations they serve.[24] Interviewees from CODI participating clinical organizations and CBOs noted that these collection instruments are in early development or are not yet widely adopted. Therefore, SDOH data are collected and represented in many ways across CODI data owners.

The Gravity Project[25] is an HL7 Fast Healthcare Interoperability Resources[26] (FHIR) accelerator project for harmonizing social risk factor data and improving electronic health information interoperability.

As part of Gravity's ongoing harmonization effort, a broad set of stakeholder groups have categorized existing SDOH data elements used for screening, diagnosis, goal setting, and interventions. They call these categories SDOH Domains.[27] For example, Gravity has categorized questions from the Hunger Vital Sign[28] screening tool to their domain "Food Insecurity" so that they can compare those questions and codes to the food insecurity questions and codes from PRAPARE,[29] another screening tool.

Gravity's Terminology Workstream Dashboard in the HL7 Confluence site has a spreadsheet for each of Gravity's SDOH domains, containing screening questions, clinical codes, and data elements that they have mapped to domains. This is work in progress and will continue to update after the major release dates of any standards that they have mapped (e.g., SNOMED: March and September, ICD: October, Logical Observation Identifiers, Names, and Codes [LOINC]: August and February).

As of early 2022 and the writing of Version 4.1 of this implementation guide, the CODI SDOH categories align with the Gravity SDOH domains. Table 7 shows a correspondence between CODI SDOH categories and Gravity SDOH domains as defined in January, 2022. The CODI categories are more general; they define topic areas without committing to any risk level. This means that the association of CODI evidence to a category does not depend on the attributed risk level. For example, any response to a screening question, whether suggesting risk or lack of risk, is still evidence. It falls to researchers to determine how best to use the available evidence.

[24] Agency for Healthcare Research and Quality, https://www.ahrq.gov/sdoh/index.html

[25] https://www.hl7.org/gravity/

[26] https://www.healthit.gov/sites/default/files/2019-08/ONCFHIRFSWhatIsFHIR.pdf

[27] https://confluence.hl7.org/display/GRAV/SDOH+Data+Elements+And+Status

[28] https://childrenshealthwatch.org/public-policy/hunger-vital-sign/

[29] Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences (PRAPARE)

Table 7: Alignment of CODI and Gravity SDOH Topic Areas

CODI SDOH Category	Definition	Gravity SDOH Domain	Definition
FOOD_DOMAIN (FD)	Pertaining to an individual's access to adequate, nutritional, safe, and culturally acceptable food.	**Food Insecurity**	Uncertain, limited, or unstable access to food that is: adequate in quantity and in nutritional quality; culturally acceptable; safe and acquired in socially acceptable ways.
HOUSING_STABILITY_DOMAIN (HS)	Pertaining to an individual's access to temporary or permanent reliable shelter.	**Housing Instability and Homelessness**	Gravity's domain definition is unavailable at the time of writing, on the Gravity Project webpage (https://confluence.hl7.org/display/GRAV/Housing+Instability+and+Homelessness).
HOUSING_ADEQUACY_DOMAIN (HA)	Pertaining to the habitability of an individual's housing.	**Inadequate Housing**	Housing does not meet habitability standards.
TRANSPORTATION_DOMAIN (TR)	Pertaining to an individual's access to transportation for routine life sustaining activities such as to place of employment, medical facilities, and school.	**Transportation Insecurity**	Uncertain, limited, or no access to safe, reliable, accessible, affordable, and socially acceptable transportation infrastructure and modalities necessary for maintaining one’s health, well-being, or livelihood.

CODI SDOH Category	Definition	Gravity SDOH Domain	Definition
INTERPERSONAL_VIOLENCE_DOMAIN (IV)	Pertaining to an individual's physical and emotional safety in close relationships.	**Intimate Partner Violence**	The term “intimate partner violence” describes physical violence, sexual violence, or psychological harm by a current or former partner or spouse. Often including a pattern of methods and tactics to gain and maintain power and control over the other person.
FINANCIAL_DOMAIN (FI)	Pertaining to an individual's ability to or feeling about meeting current and/or ongoing financial obligations.	**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.
MATERIAL_NECESSESITIES_DOMAIN (MN)	Pertaining to an individual's access to socially perceived physical necessities.	**Material Hardship**	The lack of specific socially perceived based physical necessities.
EMPLOYMENT_DOMAIN (EM)	Pertaining to an individual's status on having, looking for, or being without a job or work.	**Employment Status**	Unemployment definition: Jobless, looking for a job, and available for work.
HEALTH_INSURANCE_DOMAIN (HI)	Pertaining to an individual's access to health insurance.	----------	----------

CODI SDOH Category	Definition	Gravity SDOH Domain	Definition
ELDER_CARE_DOMAIN (EC)	Pertaining to an elder's exposure to physical, psychological, sexual, or financial abuse, or neglect by caregivers.	**Elder Abuse**	An intentional act or failure to act by a caregiver or another person in a relationship involving an expectation of trust that causes or creates a risk of harm to an older adult and can be in the form of physical abuse, psychological abuse, sexual abuse, financial abuse, and neglect by someone in a caregiving role.
EDUCATION_DOMAIN (ED)	Pertaining to an individual's academic achievements	**Educational Attainment**	Less than high education definition: Failing to meet academic criteria for high school diploma or equivalent.
VETERAN_DOMAIN (VE)	Pertaining to an individual's current and historical status in military service	**Veteran Status**	Having served as active military and honorably released or discharged.
STRESS_DOMAIN (ST)	Pertaining to an individual's perceived ability to meet, mitigate, or alter perceived excesses in environmental demands and stimuli	**Stress**	Stress: occurs when a person perceives the demands of environmental stimuli to be greater than their ability to meet, mitigate, or alter those demands.

CODI SDOH Category	Definition	Gravity SDOH Domain	Definition
SOCIAL_CONNECTION_DOMAIN (SC)	Pertaining to an individual's actual or perceived frequency of social contact, and actual or perceived access to informational, tangible, and emotional support from others.	**Social Connection**	Social Isolation: Is objectively being alone, having few relationships, or infrequent social contact. Loneliness: Is subjectively feeling alone. The discrepancy between one's desired level of connection and one's actual level. Social support: The actual or perceived availability of resources (e.g., informational, tangible, emotional) from others. Four types of social supportive behaviors: emotional, instrumental, informational and appraisal.

Appendix B Additional Guidance for CODI@NC

Organization for the North Carolina CODI Pilot

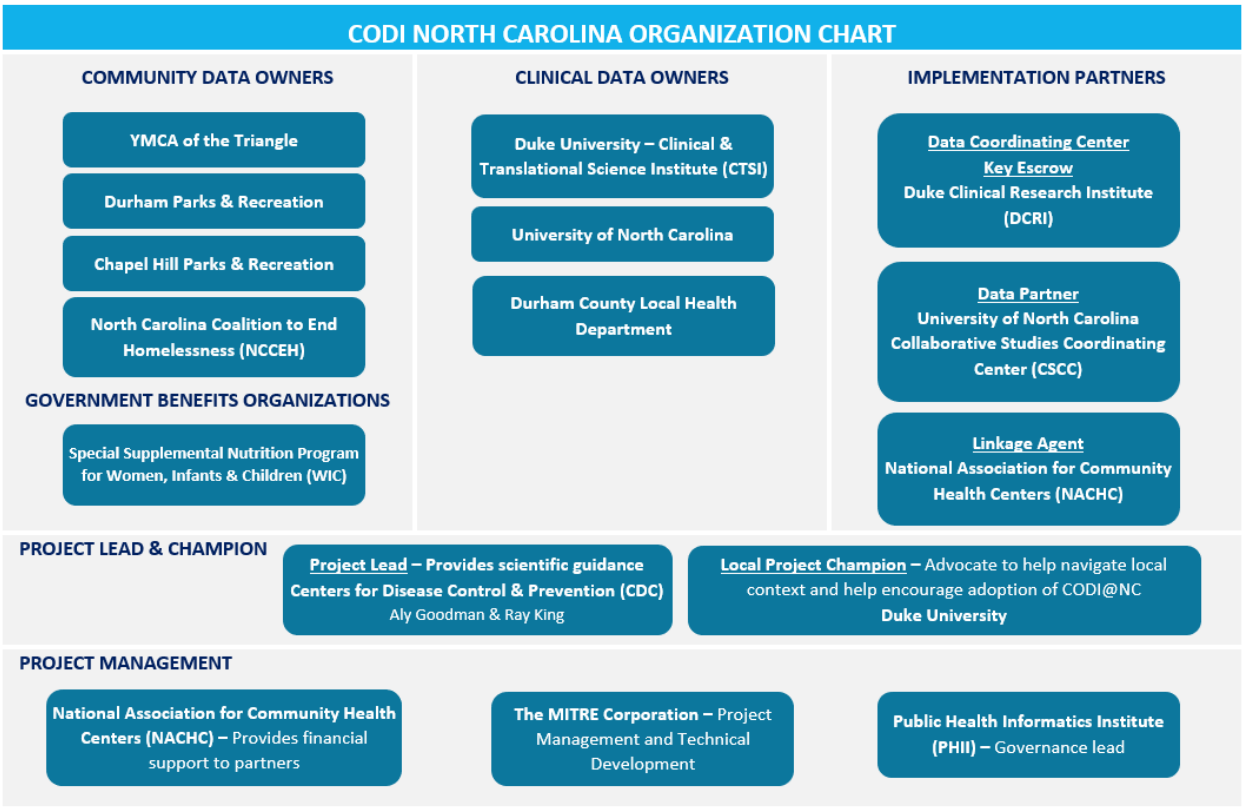


Figure 6: CODI@NC Organization Chart

Figure 6 shows the organization chart for CODI@NC (as of June 2022).

Historical Data Start Date

The start date for healthcare and program participation data populated in CODI@NC data marts was decided by the implementation work group to be January 01, 2017. Factors considered are the earliest available data from data partners, as well as the available period of population demographics data from the American Communities Survey (ACS). The RDM should be populated with information after the start date. Data partners are free to include information prior to that date if it is easier to implement the CODI tables without that date restriction or if they would prefer to make earlier data available to researchers. Researchers should be aware that data prior to the start date are necessarily incomplete.

CENSUS_LOCATION Start Date

The CHORDS VDW specification for CENSUS_LOCATION requires a LOC_START date. The default value for this field, if not otherwise known in the source data, is the date the data is extracted from the source to populate the CODI data mart.

CODI will accept multiple addresses per individual. If a data owner maintains multiple addresses per individual, then LOC_START and LOC_END values are needed for representing the chronological order of address changes. The latest known address can have a NULL LOC_END and is interpreted as the current location. Record Linkage and Data Refresh Frequency

The CODI governance committee and CCWG anticipates executing the record linkage process for the NC pilot twice per year; however, this decision may change. The CCWG expects that the data partners will perform a full data refresh after each refresh cycle of record linkage. However, alternative scenarios may be discussed with the DCC should the refresh frequency be found prohibitive for a data owner.

Physical Implementation of the CODI Data Model

The North Carolina CCWG has agreed to using the Statistical Analysis System (SAS) for the physical implementation of CODI data marts for each of the data owners. As with any DBMS, physical data types in SAS vary slightly from the logical datatypes declared in the CODI Data Model Implementation Guide (DM IG).

For the PCORnet CDM tables that CODI has adopted, the CODI data mart DLL is interpreted from the structured PCORnet CDM workbook using the RDMS data type column mapped thusly:

```
TYPE_MAPPING = {'Text' => 'varchar', 'Date' => 'date', 'Number' => 'numeric'}
```

Table 8: PCORnet CDM RDMS Type Mapping

RDBMS Types	Postgres SQL Data Type	SAS Data Type	CODI DM IG UML Data Type	Usage Notes
Date	date	SAS Date (Numeric)	date	Not to include time
Number(x)	numeric	SAS Numeric(length 8)	NUMERIC (x)	
Text(1)	N/A	SAS Char(1)	N/A	Not included in CODI's subset of PCORnet CDM
Text(10)	varchar(10)	SAS Char(10)	CHAR (10)	
Text(11)	varchar(11)	SAS Char(11)	CHAR (11)	
Text(18)	varchar (18)	SAS Char(18)	VARCHAR (18)	
Text(2)	varchar (2)	SAS Char(2)	CHAR (2)	
Text(20)	N/A	SAS Char(20)	N/A	Not included in CODI's subset of PCORnet CDM

RDBMS Types	Postgres SQL Data Type	SAS Data Type	CODI DM IG UML Data Type	Usage Notes
Text(3)	varchar (3)	SAS Char(3)	CHAR (3)	
Text(3)	varchar (3)	SAS Char(3)	LANGUAGE_TYPE	Used for language value set
Text(30)	N/A	SAS Char(30)	N/A	Not included in CODI's subset of PCORnet CDM
Text(5)	varchar (5)	SAS Char(5)	CHAR (5)	For 5 digit zip code fields
Text(5)	time	SAS Time (Numeric)	time	For time fields
Text(5)	varchar(5)	SAS Char(5)	PAYER_TYPE	For PAYER_TYPE value set
Text(8)	varchar (8)	SAS Char(8)	CHAR (8)	
Text(9)	varchar (9)	SAS Char(9)	CHAR (9)	For 9 digit zip codes
Text(x)	varchar	SAS Char(x)	CHAR (x)	

The tables below provide the mapping between the UML logical data types in this DM IG, to the physical datatypes in a Postgres SQL DMBS (used for CODI testing), and to SAS DBMS. These mappings address all the CODI ancillary table fields, and some PCORnet CDM table fields. For a complete mapping of PCORnet CDM data type mapping to SAS data types refer to the PCORnet CDM.

Table 9 covers the mapping between primary UML data types declared in the DM IG to recommended primary types in Postgres SQL and SAS databases. This table includes the CODI data type, ID.

Table 9: Ancillary CODI DM Data type Mapping

CODI DM IG UML Data Type	Postgres SQL Data Type	SAS Data Type	RDBMS Types	Usage Notes
Boolean	Boolean	SAS Numeric (length 1)	Number (1)	
CHAR (10)	char (10)	SAS Char(10)	Text (10)	Phone number
CHAR (5)	char (5)	SAS Char(5)	Text (5)	5 digit zip code
CHAR (9)	char (9)	SAS Char(9)	Text (9)	9 digit zip code
date	date	SAS Date (Numeric)	Date	Not to include time
float	float	SAS Numeric (length 8)	Number (x)	

CODI DM IG UML Data Type	Postgres SQL Data Type	SAS Data Type	RDBMS Types	Usage Notes
ID	varchar	SAS Char(x)	Text (x)	For all ID primary key fields and ID foreign key fields
Integer	integer	SAS Numeric (length 8)	Number (x)	
NUMERIC(8, 6)	decimal (8, 6)	SAS Numeric(length 8)	Number (x)	Latitude
NUMERIC(9, 6)	decimal (9, 6)	SAS Numeric(length 8)	Number (x)	Longitude
NUMERIC(x)	numeric	SAS Numeric (length 8)	Number (x)	Census boundary year (as defined by CHORDS VDW)
String	varchar	SAS Char(x)	Text (x)	For description, explanation, and uncoded reason fields.
String	varchar (255)	SAS Char(x)	Text (255)	For name and address fields without a value set constraint.
time	time	SAS Time (Numeric)	Text (5)	
VARCHAR (10)	varchar(10)	SAS Char(10)	Text (10)	For LOINC codes
VARCHAR (15)	varchar(15)	SAS Char(15)	Text (15)	Geocode
VARCHAR (18)	varchar (18)	SAS Char(18)	Text (18)	Condition codes

The CODI DM IG data dictionary treats a value set (aka, codeset) as a unified modeling language (UML) enumeration type, which is a user-defined data type. Ultimately, in a relational database, these enumerated data types are converted to a primary data type, with a check constraint on a set of allowable values. Refer to Table 10 for named value sets and their corresponding primary data types.

Table 10 contains only those value sets used by fields in tables that are owned by CODI. Some value sets are defined specifically for CODI while others are reused from PCORnet or other standards. To find the physical data type for fields with value sets in tables defined by PCORnet, refer to the PCORnet CDM specification.

Table 10: Primary Data Type for Named Value Sets

Value Set Owner	CODI DM IG Value Set Data Type	Postgres SQL Data Type	RDBMS Type	SAS Data Type
CODI	ASSET_TYPE	char (2)	Text (2)	SAS Char(2)
CODI	DIRECTION_TYPE	char (1)	Text (1)	SAS Char(1)
PCORnet CDM	DX_TYPE	char (2)	Text (2)	SAS Char(2)
CODI	FREQ_TYPE	char (1)	Text (1)	SAS Char(1)
VDW	GEOLEVEL_TYPE	char (1)	Text (1)	SAS Char(1)
CODI	MODE_TYPE	char (1)	Text (1)	SAS Char(1)
CMS Certification Number (CCN)	ORGANIZATION_TYPE	varchar (6)	Text (6)	SAS Char(6)
CODI	PROCESS_PERFORMED_TYPE	char (2)	Text (2)	SAS Char(2)
CODI	REFERRAL_STATUS_TYPE	char (2)	Text (2)	SAS Char(2)
CODI	RELATIONSHIP_TYPE	varchar (9)	Text (9)	SAS Char (9)
CODI	SDOH_CATEGORY_TYPE	char (2)	Text (2)	SAS Char(2)
CODI	SETTING_TYPE	char (2)	Text (2)	SAS Char(2)
PCORnet CDM	SPECIALTY_TYPE	varchar (10)	Text (x)	SAS Char(x)

Appendix C CODI Record Linkage Data Model Documentation

The CODI Record Linkage Data Model uses two PCORnet CDM tables and two CODI ancillary tables to support the privacy preserving record linkage (PPRL) process. Record linkage enables the creation of individual-level longitudinal records.

PRIVATE_DEMOGRAPHIC AND PRIVATE_ADDRESS_HISTORY, from the PCORnet CDM, contain PII. Therefore, the PPRL IG recommends that sites establish sufficient business rules to protect these sensitive data, such as housing them separately from research data and deleting the private data once the linkage process is complete.

The two CODI ancillary tables, LINK and HOUSEHOLD_LINK, contain the result of the privacy preserving record linkage process. LINK contains a unique, but anonymous ID for each patient or program participant used to link that individual's records across CODI data providers. HOUSEHOLD_LINK contains an anonymous household ID, linking individuals living in the same household to each other. These two link tables are incorporated into the RDM to support clinical and community data queries.

The data dictionary for the RLDM follows.

HOUSEHOLD_LINK

A household link represents a connection between people (identified anonymously) who have the same physical address at the time the household link is established. Household linkage is not based on family relationships.

The HOUSEHOLD_LINK table contains one record for each person in the demographics table for each iteration of a household record linkage. Each iteration establishes new household IDs for households, and therefore cannot provide longitudinal household membership information.

HOUSEHOLDID, PATID, and LINK_ITERATION make up the composite primary key for HOUSEHOLD_LINK. This primary key pattern is distinct from the PCORnet (and most of CODI) primary key pattern in which every table has a single technical key whose name is [TABLE_NAME]_ID or [TABLENAME]ID.

Table 11. HOUSEHOLD_LINK Details

Attribute	Cardinality	Type	Documentation
PATID	1	FK::DEMOGRAPHIC	A member of a household who is a person and who is anonymously identified in the demographics table.
HOUSEHOLDID	1	ID	An identifier for a (virtual) household that is unique across CODI data marts in the same CODI network. More than one patient or participant (PATID) may be linked to the same household (HOUSEHOLDID)
LINK_ITERATION	1	Integer	An iteration of the household record linkage process.

LINK

The LINK table contains one record for each person in the demographics table for each iteration of record linkage. Each iteration establishes a new LINKID for each person.

Table 12. LINK Details

Attribute	Cardinality	Type	Documentation
PATID	1	FK::DEMOGRAPHIC	A link back to the demographics table.

Attribute	Cardinality	Type	Documentation
LINKID	1	ID	A globally unique identifier linking individuals across data sources.
LINK_ITERATION	1	Integer	An iteration of the record linkage process.

PRIVATE_ADDRESS_HISTORY

[From PCORnet CDM]

Protected table that can be used to store elements of a patient's address that are considered personal health information (PHI). These data can be used for geocoding or other linkage projects.

Table 13. PRIVATE_ADDRESS_HISTORY Details

Attribute	Cardinality	Type	Documentation
ADDRESSID	1	ID	Arbitrary identifier for each unique address record.
PATID	1	FK::PRIVATE_DEMOGRAPHIC	Arbitrary person-level identifier. Used to link across tables.
ADDRESS_STREET	0..1	String	Primary address line (e.g., street name and number)
ADDRESS_DETAIL	0..1	String	Remaining address details (e.g., suite, post office box, other details)
ADDRESS_CITY	0..1	String	The name of the city, town, village or other community
ADDRESS_ZIP5	0..1	CHAR (5)	5-digit postal code for the address.
ADDRESS_STATE	0..1	STATE	State, as represented by 2-digit postal abbreviation.
ADDRESS_TYPE	1	ADDRESS_TYPE_TYPE	Type of address. Details of categorical definitions: Postal: mailing address – PO Boxes and care of addresses. Physical: A physical address that can be visited. Both: An address that is both physical and postal.
ADDRESS_PREFERRED	1	YES//NO	Indicates whether this address is the preferred one for a given patient, address use and address type within a given address period.

Attribute	Cardinality	Type	Documentation
ADDRESS_USE	1	ADDRESS_USE_TYPE	Purpose of the address. Details of categorical definitions: Home: A communication address at home. Work: An office address. First choice for business-related contacts during business hours. Temp: A temporary address. Old/Incorrect: This address is no longer in use (or was never correct but retained for records).
ADDRESS_ZIP9	0..1	CHAR(9)	9-digit postal code for the address.
ADDRESS_PERIOD_START	0..1	Date	Initial date when the address known to be in use.
ADDRESS_PERIOD_END	0..1	Date	Date when address was no longer in use.
RAW_ADDRESS_TEXT	0..1	String	Text representation of the address

PRIVATE_DEMOGRAPHIC

[From PCORnet CDM]

Protected table that is intended to provide a standardized representation of the personally-identifiable information (PII) that is needed to support local activities related to record linkage.

Table 14. PRIVATE_DEMOGRAPHIC Details

Attribute	Cardinality	Type	Documentation
PATID	1	ID	Arbitrary person-level identifier. Used to link across tables. Corresponds to PATID in the DEMOGRAPHIC table
PAT_FIRSTNAME	0..1	String	Given name of the patient.
PAT_MIDDLENAME	0..1	String	Middle name of the patient.
PAT_LASTNAME	0..1	String	Surname or family name.
PAT_MAIDENNAME	0..1	String	Surname or family name prior to marriage.
BIRTH_DATE	0..1	date	Date of birth. Corresponds to BIRTH_DATE in the DEMOGRAPHIC table.

Attribute	Cardinality	Type	Documentation
SEX	0..1	SEX_TYPE	Sex assigned at birth. Corresponds to SEX in the DEMOGRAPHIC table.
RACE	0..1	RACE_TYPE	Please use only one race value per patient. Corresponds to RACE in the DEMOGRAPHIC table.
HISPANIC	0..1	YES//NO	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Corresponds to HISPANIC in the DEMOGRAPHIC table.
PRIMARY_EMAIL	0..1	String	Primary e-mail address for the patient.
PRIMARY_PHONE	0..1	CHAR (10)	Primary phone number for the patient (if known). 10-digit US phone number.
PrivateAddressHistory	0..*	FK::PRIVATE_ADDRESS_HISTORY	An address for this patient used for record linkage at the household level and geo-coding.

Appendix D CODI Research Data Model Dictionary

Overview

The CODI Research Data Model (RDM) incorporates tables from three data model specifications and adds a number of tables designed specifically for CODI. CODI's RDM is primarily based on a subset of the PCORnet CDM version 6.0. The RDM also adopts tables from CHORDS VDW version 3.5 and from OMOP version 6.0. The borrowed tables are depicted in Figure 7. The rest of the tables in the RDM are unique to CODI and are shown in Figure 8.

Notation

Within Figure 7 and Figure 8 color distinguishes the three RDM sections.

- White tables are from PCORnet CDM.
- Yellow tables from CHORDS VDW or OMOP.
- Orange tables are data tables unique to CODI.

Other notation and label formats within these figures and in the data dictionary, are intentionally used to convey the following information.

- Lines between table boxes indicate PK/FK relationships among the tables.
- Labels on the ends of lines that are in ALL CAPS are the foreign key column names (e.g., PATID is a is a foreign key column in the ENCOUNTER table referencing the DEMOGRAPHIC table).
- Numbers on the ends of lines indicate cardinality constraints.
  - 0..1 means a column is optional.
  - 1 means a column is required.
  - 0..* and 1..* document the intended meaning of a relationship; these will not appear in a CODI warehouse (they are the inverses of FK columns).
- Labels in CamelCase are provided to document the intended meaning of the relationship; these will not appear in a CODI data warehouse.
- Labels in lower case are unique to OMOP due to a difference in notation standards and can be interpreted the same as those in ALL CAPS.

Figure 7. CODI Research Data Model Tables from PCORnet CDM, VDW, and OMOP

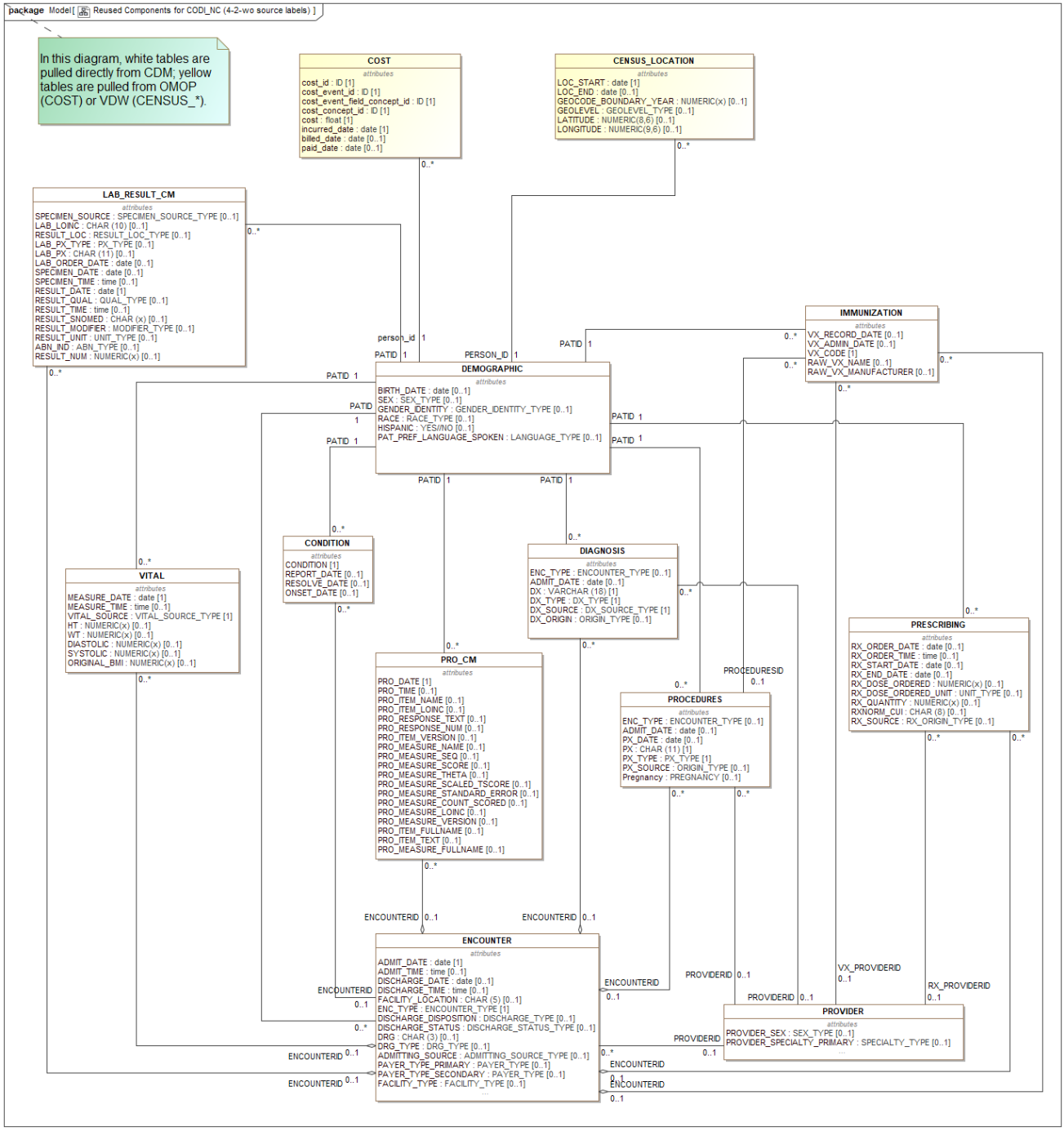
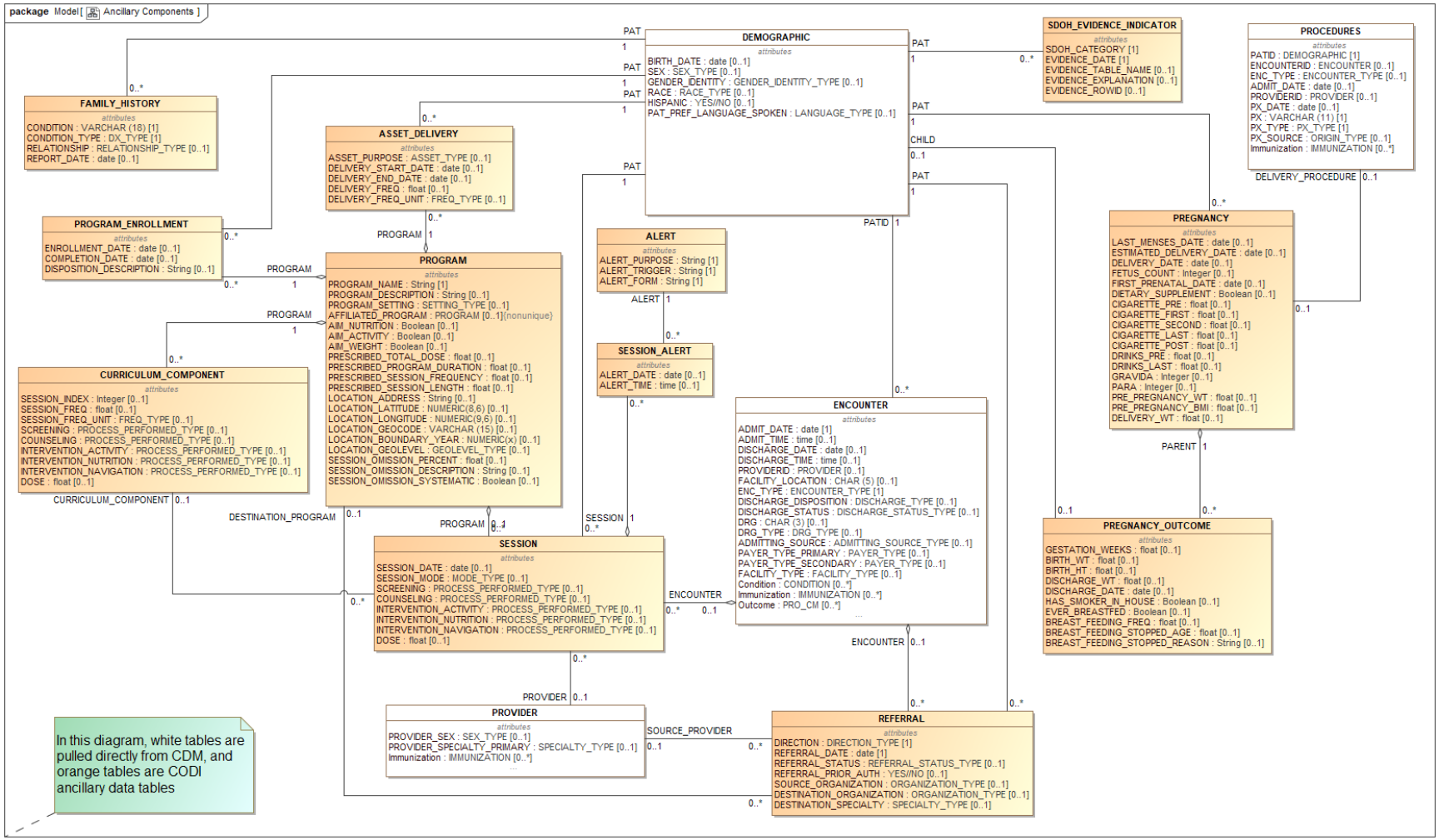


Figure 8. CODI Research Data Model Ancillary Tables



CODI Research Data Model Table Details

The tables and attributes listed in this section are either from the CODI ancillary portion of the RDM, or are borrowed from PCORnet CDM 6.0, CHORDS VDW 3.5, or OMOP 6.0. For the borrowed tables, this document is not the source of record nor the complete implementation guidance. Borrowed tables are listed here for convenience and to declare the specific subset of those other models on which CODI relies. The table definition will indicate when a table is borrowed from another data model specification.

ALERT

The ALERT table contains one record for each distinct kind of alert. Alerts are components of a clinical decision support system (CDS). Given the gamut of possible alerts and the idiosyncrasies of CDS implementations, CODI only captures a prose description of the intended function of the alert. Only obesity- or weight-related alerts should be captured for CODI.

Table 15. ALERT Details

Attribute	Cardinality	Type	Documentation
ALERTID	1	ID	
ALERT_PURPOSE	1	String	A description of the purpose of the alert.
ALERT_TRIGGER	1	String	A description of the conditions under which the alert triggers.
ALERT_FORM	1	String	A description of how the alert is presented to the user.
SessionAlert	0..*	FK::SESSION_ALERT	A set of sessions during which this alert triggered.

ASSET_DELIVERY

The ASSET_DELIVERY table contains one record for each contiguous period of time during which a person consistently receives assets. An asset is a resource transferred by a program to an individual.

Table 16. ASSET_DELIVERY Details

Attribute	Cardinality	Type	Documentation
ASSET_DELIVERY_ID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographic table.

Attribute	Cardinality	Type	Documentation
PROGRAMID	1	FK::PROGRAM	A link back to the program table.
ASSET_PURPOSE	0..1	ASSET_TYPE	An intended purpose for the use of a monetary asset (e.g., health insurance or food).
DELIVERY_START_DATE	0..1	date	A date the asset delivery began.
DELIVERY_END_DATE	0..1	date	A date the asset delivery ended.
DELIVERY_FREQ	0..1	float	A number of times an asset is delivered each unit of time.
DELIVERY_FREQ_UNIT	0..1	FREQ_TYPE	A unit of time used to describe how often an asset is delivered. For example, an asset delivered twice a week has a frequency of 2 and a unit of Weekly. An asset delivered every other week has a frequency of 0.5 and a unit of Weekly.

CENSUS_LOCATION

[From CHORDS VDW]

The CENSUS_LOCATION table holds information about the patient's location based on their address stored as a geocode. To complete this table, and populate the related geo-demographic data, patient addresses must be geocoded. Every PERSON_ID appearing in the demographics table must have a record in this table.

Detailed instructions are in the CENSUS_LOCATION Section 4.2.3 of this CODI IG and in the CHORDS VDW IG.

Table 17. CENSUS_LOCATION Details

Attribute	Cardinality	Type	Documentation
PERSON_ID	1	FK::DEMOGRAPHIC	Person ID.
LOC_START	1	date	Beginning date address for individual is valid (if date is unknown, assign a date estimated to when electronic address data began collection.
LOC_END	0..1	date	Ending date address for individual is valid (dates prior to 01/01/2010 should default to 01/01/2010. If location record is a patient's current location, leave null. Queries will use this criterion to identify a patient's current address.

Attribute	Cardinality	Type	Documentation
GEOCODE	0..1	varchar (15)	Numeric string of multiple census variables. The geocode can be up to 15 digits long with geographic information hierarchically documented through the string of digits: state(2) + county(3) + tract(6), blockgp(1) + block(3) = GeoCode(15) When an address can be geocoded to a tract, the geocode is 11 digits long. When an address cannot be geocoded to a tract but can be geocoded to a county, the geocode is 5 digits long. Geocodes with block group and block information should be truncated to 11 digits to meet the requirements of a limited dataset (state(2) + county(3) + tract(6)).
GEOCODE_BOUNDARY_YEAR	0..1	NUMERIC(8)	Indicating the census year for which geocode applies (year in which geocode is valid)
GEOLEVEL	0..1	GEOLEVEL_TYPE	Indicates the specificity of the GEOCODE match This can be assessed using logic that considers the length of the GEOCODE value (2 characters for state; 5 characters for county; 11 characters for census tract)Data partners should have the values T or C. Data partners should not map to B or G.
LATITUDE	0..1	NUMERIC(8,6)	The latitude of the location.
LONGITUDE	0..1	NUMERIC(9,6)	The longitude of the location.

CONDITION

[From PCORnet CDM]

A condition represents a patient's diagnosed and self-reported health conditions and diseases. The patient's medical history and current state may both be represented.

Table 18. CONDITION Details

Attribute	Cardinality	Type	Documentation
CONDITIONID	1	ID	
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTER	Arbitrary encounter-level identifier used to link across tables. This is an optional field and should only be populated if the item was collected as part of a

Attribute	Cardinality	Type	Documentation
			healthcare encounter. If more than one encounter association is present, this field should be populated with the ID of the encounter when the condition was first entered into the system. However, please note that many conditions may be recorded outside of an encounter context.
CONDITION	1	VARCHAR (18)	Some codes will contain leading zeroes, and different levels of decimal precision may also be present. This field is a character field, not numeric, to accommodate these coding conventions. Please populate the exact value of this diagnosis code, but remove any source-specific suffixes and prefixes. (Description updated in v3.1.)
CONDITION_SOURCE	1	CONDITION_SOURCE_TYPE	Please note: The “Patient-reported” category can include reporting by a proxy, such as patient’s family or guardian.
CONDITION_STATUS	0..1	CONDITION_STATUS_TYPE	Condition status corresponding with REPORT_DATE.
CONDITION_TYPE	1	CONDITION_TYPE_TYPE	Condition code type. Please note: The “Other” category is meant to identify internal use ontologies and codes.
REPORT_DATE	0..1	date	Date condition was noted, which may be the date when it was recorded by a provider or nurse, or the date on which the patient reported it. Please note that this date may not correspond to onset date.
RESOLVE_DATE	0..1	date	Date condition was resolved, if resolution of a transient condition has been achieved. A resolution date is not generally expected for chronic conditions, even if the condition is managed.
ONSET_DATE	0..1	date	The onset date concept here refers to "the date and time when problem (illness, disorder, or symptom) started" (ONC:MU Clinical Data Set, caDSR 4973971). This is a different concept than report date, which is the date on which the medical status was collected. An onset date should generally be considered independently of the observer or provider. However, the judgment of when a condition "started" depends on the disease, the frequency of visits, and many other factors. It is not clear that any facility or physician employs this field in a manner which can be trusted without validation during analysis. (New definition added in v3.1.)

COST

[From OMOP]

The COST table captures records containing the cost of any medical event recorded in one of the OMOP clinical event table.

1) This table does not capture the cost of providing the service, but rather the amounts billed and received.

2) The COST table can link to ENCOUNTER, LAB_RESULT_CM, PROCEDURES, or SESSION.

Table 19. COST Details

Attribute	Cardinality	Type	Documentation
cost_id	1	ID	A unique identifier for each COST record.
person_id	1	FK::DEMOGRAPHIC	A unique identifier for each PERSON.
cost_event_id	1	ID	A foreign key identifier to the event (e.g. Measurement, Procedure, Visit, Drug Exposure, etc) record for which cost data are recorded. 1) A reference to one of ENCOUNTER, LAB_RESULT_CM, PROCEDURES, or SESSION.
cost_event_field_concept_id	1	ID	A foreign key identifier to a concept in the CONCEPT table representing the identity of the field represented by COST_EVENT_ID 1) One of ENCOUNTER, LAB_RESULT_CM, PROCEDURES, or SESSION.
cost_concept_id	1	ID	A foreign key that refers to a Standard Cost Concept identifier in the Standardized Vocabularies belonging to the 'Cost' vocabulary. 1) CODI will only capture charge data, not payments.
cost	1	float	The actual financial cost amount
incurred_date	1	date	The first date of service of the clinical event corresponding to the cost as in table capturing the information (e.g. date of visit, date of procedure, date of condition, date of drug etc).
billed_date	0..1	date	The date a bill was generated for a service or encounter
paid_date	0..1	date	The date payment was received for a service or encounter

CURRICULUM_COMPONENT

A curriculum component is a standard element of a program. A program can comprise a fixed curriculum with a predefined endpoint and an enumerated set of standard sessions. Alternatively, a program can comprise a recurring curriculum with no endpoint and a set of standard sessions that recur with some frequency.

Table 20. CURRICULUM_COMPONENT Details

Attribute	Cardinality	Type	Documentation
CURRICULUM_COMPONENT_ID	1	ID	
PROGRAMID	1	FK::PROGRAM	A link back to the program this component of a curriculum belongs to.
SESSION_INDEX	0..1	integer	An ordinal used to establish a total ordering on the sessions within a fixed curriculum.
SESSION_FREQ	0..1	float	A number of times a session is administered each unit of time.
SESSION_FREQ_UNIT	0..1	FREQ_TYPE	A unit of time used to describe how often a session is administered. For example, a session administered twice a week has a frequency of 2 and a unit of Weekly. A session administered every other week has a frequency of 0.5 and a unit of Weekly.
SCREENING	0..1	PROCESS_PERFORMED_TYPE	True if the sessions associated with this curriculum include any assessment of lifestyle behaviors related to obesity, such as physical activity, nutrition, screen time, or sleep.
COUNSELING	0..1	PROCESS_PERFORMED_TYPE	True if the sessions associated with this curriculum include any advice or direction regarding lifestyle related to obesity, such as physical activity, nutrition, screen time, or sleep.
INTERVENTION_ACTIVITY	0..1	PROCESS_PERFORMED_TYPE	True if the sessions associated with this curriculum include performing at least moderate physical activity; moderate activity requires a moderate amount of effort (5-6 on a scale of 0 to 10) and noticeably accelerates the heart rate and breathing.
INTERVENTION_NUTRITION	0..1	PROCESS_PERFORMED_TYPE	True if the sessions associated with this curriculum include an activity designed to improve nutrition.
INTERVENTION_NAVIGATION	0..1	PROCESS_PERFORMED_TYPE	True if the sessions associated with this curriculum include a navigational service to access benefits or to overcome barriers to care.

Attribute	Cardinality	Type	Documentation
DOSE	0..1	float	A measure of the amount of time sessions associated with this curriculum are expected to last.
Session	0..*	FK::SESSION	A set of sessions associated with this curriculum component.

DEMOGRAPHIC

[From PCORnet CDM]

The DEMOGRAPHIC table contains a single record for each patient. Demographics record the direct attributes of individual patients.

Table 21. DEMOGRAPHIC Details

Attribute	Cardinality	Type	Documentation
PATID	1	ID	
BIRTH_DATE	0..1	date	Date of birth.
SEX	0..1	SEX_TYPE	Sex assigned at birth.
GENDER_IDENTITY	0..1	GENDER_IDENTITY_TYPE	Current gender identity.
HISPANIC	0..1	YES//NO	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race.
RACE	0..1	RACE_TYPE	Please use only one race value per patient.
PAT_PREF_LANGUAGE_SPOKEN	0..1	LANGUAGE_TYPE	Preferred spoken language of communication as expressed by the patient.
Birth	0..1	FK::PREGNANCY_OUTCOME	A pregnancy outcome for this individual describing this individual's birth.
CensusLocation	0..*	FK::CENSUS_LOCATION	A set of census locations associated with this person.
Condition	0..*	FK::CONDITION	A set of conditions associated with this person.
Cost	0..*	FK::COST	A set of charges associated with this person.
FamilyHistory	0..*	FK::FAMILY_HISTORY	A set of family histories associated with this person.

Attribute	Cardinality	Type	Documentation
Encounter	0..*	FK::ENCOUNTER	A set of encounters associated with this person.
AssetDelivery	0..*	FK::ASSET_DELIVERY	A set of asset deliveries associated with this person.
ProgramEnrollment	0..*	FK::PROGRAM_ENROLLMENT	A set of enrollments associated with this person.
Diagnosis	0..*	FK::DIAGNOSIS	A set of diagnoses associated with this person.
Immunization	0..*	FK::IMMUNIZATION	A set of immunizations associated with this person.
LabResult	0..*	FK::LAB_RESULT_CM	A set of lab results associated with this person.
Outcome	0..*	FK::PRO_CM	A set of patient-reported outcomes associated with this person.
Pregnancy	0..*	FK::PREGNANCY	A set of pregnancies associated with this person.
Prescribing	0..*	FK::PRESCRIBING	A set of prescriptions associated with this person.
Procedure	0..*	FK::PROCEDURES	A set of procedures associated with this person.
SdohEvidence	0..*	FK::SDOH_EVIDENCE_INDICATOR	A set of SDOH evidence indicators associated with this person.
Session	0..*	FK::SESSION	A set of sessions associated with this person.
Referral	0..*	FK::REFERRAL	A set of referrals associated with this person.
Vital	0..*	FK::VITAL	A set of vital signs associated with this person.
HouseholdLink	0..1	FK::HOUSEHOLD_LINK	A set of household linkage identifiers associated with this person.
Link	0..1	FK::LINK	A set of record linkage identifiers associated with this person.

DIAGNOSIS

[From PCORnet CDM]

The DIAGNOSIS table contains one record for each diagnosis of a patient. Diagnosis codes indicate the results of diagnostic processes and medical coding within healthcare delivery. Data in this table are expected to be from healthcare-mediated processes and reimbursement drivers.

Table 22. DIAGNOSIS Details

Attribute	Cardinality	Type	Documentation
DIAGNOSISID	1	ID	
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTER	Arbitrary encounter-level identifier. Used to link across tables.
ENC_TYPE	0..1	ENCOUNTER_TYPE	This is a field replicated from the ENCOUNTER table. See the ENCOUNTER table for definitions.
ADMIT_DATE	0..1	date	This is a field replicated from the ENCOUNTER table. See the ENCOUNTER table for definitions.
PROVIDERID	0..1	FK::PROVIDER	Identifier associated with the provider most responsible for the diagnosis.
DX	1	VARCHAR (18)	Diagnosis code.
DX_TYPE	1	DX_TYPE	Diagnosis code type. We provide values for ICD and SNOMED code types. Other code types will be added as new terminologies are more widely used.
DX_SOURCE	1	DX_SOURCE_TYPE	Classification of diagnosis source. We include these categories to allow some flexibility in implementation. The context is to capture available diagnoses recorded during a specific encounter.
DX_ORIGIN	0..1	ORIGIN_TYPE	Source of the diagnosis information.

ENCOUNTER

[From PCORnet CDM]

The ENCOUNTER table contains one record for each unique encounter. Encounters are interactions between patients and providers within the context of healthcare delivery. An encounter comprises multiple visits, diagnoses, procedures, etc.

Table 23. ENCOUNTER Details

Attribute	Cardinality	Type	Documentation
ENCOUNTERID	1	ID	
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier used to link across tables.
ADMIT_DATE	1	date	Encounter or admission date.
ADMIT_TIME	0..1	time	Encounter or admission time.
DISCHARGE_DATE	0..1	date	Discharge date.
DISCHARGE_TIME	0..1	time	Discharge time.
PROVIDERID	0..1	FK::PROVIDER	Code for the provider who is most responsible for this encounter. As with the PATID, the provider code is a pseudoidentifier with a consistent crosswalk to the real identifier.
FACILITY_LOCATION	0..1	CHAR (5)	Geographic location (5 digit zip code).
ENC_TYPE	1	ENCOUNTER_TYPE	Encounter type.
DISCHARGE_DISPOSITION	0..1	DISCHARGE_TYPE	Vital status at discharge.
DISCHARGE_STATUS	0..1	DISCHARGE_STATUS_TYPE	Discharge status.
DRG	0..1	CHAR (3)	3-digit Diagnosis Related Group (DRG).
DRG_TYPE	0..1	DRG_TYPE	DRG code version.
ADMITTING_SOURCE	0..1	ADMITTING_SOURCE_TYPE	Admitting source.
PAYER_TYPE_PRIMARY	0..1	PAYER_TYPE	Categorization of payer type for primary payer associated with the encounter.
PAYER_TYPE_SECONDARY	0..1	PAYER_TYPE	Categorization of payer type for secondary payer associated with the encounter.
FACILITY_TYPE	0..1	FACILITY_TYPE	Description of the facility where the encounter occurred.
Condition	0..*	FK::CONDITION	A set of conditions associated with this encounter.
Diagnosis	0..*	FK::DIAGNOSIS	A set of diagnoses associated with this encounter.
Immunization	0..*	FK::IMMUNIZATION	A set of immunizations associated with this encounter.

Attribute	Cardinality	Type	Documentation
LabResult	0..*	FK::LAB_RESULT_CM	A set of lab results associated with this encounter.
Outcome	0..*	FK::PRO_CM	A set of patient-reported outcomes associated with this encounter.
Prescribing	0..*	FK::PRESCRIBING	A set of prescriptions associated with this encounter.
Procedure	0..*	FK::PROCEDURES	A set of procedures associated with this encounter.
Referral	0..*	FK::REFERRAL	A set of referrals associated with this encounter.
Session	0..*	FK::SESSION	A set of sessions associated with this encounter.
Vital	0..*	FK::VITAL	A set of vital signs associated with this encounter.

FAMILY_HISTORY

The FAMILY_HISTORY table stores information regarding an individual's family history of disease. A separate record is created for each report of a condition that a family member has. Absence of a record in this table is not indicative the absence of a condition.

This information is intended to be pulled from the patient's record, not by linking to a family member's medical record.

Table 24. FAMILY_HISTORY Details

Attribute	Cardinality	Type	Documentation
FAMILY_HISTORY_ID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographic table.
CONDITION	1	VARCHAR (18)	A condition that the patient has a family history of.
CONDITION_TYPE	1	DX_TYPE	A condition coding system from which the condition code is drawn.
RELATIONSHIP	0..1	RELATIONSHIP_TYPE	An indication of which relative has the condition
REPORT_DATE	0..1	date	A date the family history of the condition was reported.

HARVEST

[From PCORnet CDM]

Attributes associated with the specific PCORnet [or CODI] datamart implementation, including data refreshes.

Table 25. HARVEST Details

Attribute	Cardinality	Type	Documentation
HARVESTID	1	ID	
CDM_VERSION	0..1	CHAR (3)	Version currently implemented within this datamart.
DATAMART_EHR	0..1	DATAMART_EHR_TYPE	Datamart includes EHR data source(s)
DATAMART_NAME	0..1	String	Descriptive name of the datamart. This identifier is assigned by the PCORnet Distributed Research Network Operations Center (DRN OC)
DATAMART_PLATFORM	0..1	DATAMART_PLATFORM_TYPE	No documentation provided by PCORnet CDM. CODI guidance: use for identifying the datamart database management system type.
DATAMARTID	0..1	VARCHAR(10)	This identifier is assigned by the PCORnet Distributed Research Network Operations Center (DRN OC)
REFRESH_DEMOGRAPHIC_DATE	0..1	date	Most recent date on which the present data were loaded into the DEMOGRAPHIC table. This date should be null if the table does not have records. CODI Guidance: This date applies to all tables in the datamart, because CODI will refresh all tables together.

IMMUNIZATION

[From PCORnet CDM]

The IMMUNIZATION table contains records of immunizations that have been delivered within the health system as well as reports of those administered elsewhere.

Table 26. IMMUNIZATION Details

Attribute	Cardinality	Type	Documentation
IMMUNIZATIONID	1	ID	
VX_RECORD_DATE	0..1	date	Date immunization was recorded (i.e., date record was created).
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier used to link across tables.

Attribute	Cardinality	Type	Documentation
ENCOUNTERID	0..1	FK::ENCOUNTER	Arbitrary encounter-level identifier. This should be present if the immunization activity is directly associated with an encounter.
PROCEDURESID	0..1	FK::PROCEDURES	This is an optional relationship to the PROCEDURES table and is not expected to be available for all immunizations. One procedure may generate multiple immunization records.
VX_PROVIDERID	0..1	FK::PROVIDER	Provider code for the provider who delivered the immunization. The provider code is a pseudo-identifier with a consistent crosswalk to the real identifier.
VX_ADMIN_DATE	0..1	date	Date immunization was administered, if known.
VX_CODE_TYPE	1	VX_CODE_TYPE_TYPE	Immunization code type.
VX_CODE	1	String	Immunization code
VX_STATUS	1	VX_STATUS_TYPE	Status of the immunization.
VX_STATUS_REASON	0..1	VX_STATUS_REASON_TYPE	Reason immunization is incomplete or not done.
VX_SOURCE	0..1	VX_SOURCE_TYPE	Source of the prescribing information.
VX_MANUFACTURER	0..1	VX_MANUFACTURER_TYPE	Manufacturer of the immunization.
RAW_VX_NAME	0..1	String	Field for originating, full textual immunization name from the source.
RAW_VX_MANUFACTURER	0..1	String	Field for originating value, prior to mapping into the PCORnet CDM value set.

LAB_RESULT_CM

[From PCORnet CDM]

The LAB_RESULT_CM table contains one record for each lab result of a patient. This table is used to store quantitative and qualitative measurements from blood and other body specimens.

Table 27. LAB_RESULT_CM Details

Attribute	Cardinality	Type	Documentation
LAB_RESULT_CM_ID	1	ID	
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTER	Arbitrary encounter-level identifier. Not all lab results will be associated with a healthcare encounter.
SPECIMEN_SOURCE	0..1	SPECIMEN_SOURCE_TYPE	Specimen source. All records will have a specimen source; some tests have several possible values for SPECIMEN_SOURCE.
LAB_LOINC	0..1	VARCHAR (10)	Use this field to store the LOINC code of the laboratory result.
RESULT_LOC	0..1	RESULT_LOC_TYPE	Location of the test result. Point of Care locations may include anticoagulation clinic, newborn nursery, finger stick in provider office, or home. The default value is 'L' unless the result is Point of Care. There should not be any missing values.
LAB_PX_TYPE	0..1	PX_TYPE	Procedure code type, if applicable.
LAB_PX	0..1	VARCHAR (11)	Variable for local and standard procedure codes, used to identify the originating order for the lab test.
LAB_ORDER_DATE	0..1	date	Date test was ordered.
SPECIMEN_DATE	0..1	date	Date specimen was collected.
SPECIMEN_TIME	0..1	time	Time specimen was collected.
RESULT_DATE	1	date	Result date.
RESULT_QUAL	0..1	QUAL_TYPE	Standardized result for qualitative results. This variable should be NI for quantitative results.
RESULT_TIME	0..1	time	Result time.
RESULT_SNOMED	0..1	VARCHAR	If the qualitative result has been mapped to SNOMED CT, the corresponding SNOMED code can be placed here.
RESULT_MODIFIER	0..1	MODIFIER_TYPE	Modifier for result values.
RESULT_UNIT	0..1	UNIT_TYPE	Converted/standardized units for the quantitative result.
ABN_IND	0..1	ABN_TYPE	Abnormal result indicator. This value comes from the source data; do not apply logic to create it.
RESULT_NUM	0..1	NUMERIC(x)	Standardized/converted result for quantitative results.

PREGNANCY

The PREGNANCY table contains one record for each pregnancy.

Table 28. PREGNANCY Details

Attribute	Cardinality	Type	Documentation
PREGNANCYID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographics table.
LAST_MENSES_DATE	0..1	date	A date of the parent's last menstrual period.
ESTIMATED_DELIVERY_DATE	0..1	date	An estimated date of delivery.
DELIVERY_DATE	0..1	date	An actual date of delivery.
FETUS_COUNT	0..1	Integer	A number of fetuses involved in this pregnancy.
FIRST_PRENATAL_DATE	0..1	date	A date of the parent's first prenatal healthcare visit.
DIETARY_SUPPLEMENT	0..1	Boolean	True if the parent took dietary supplements during pregnancy.
CIGARETTE_PRE	0..1	float	A number of cigarettes the parent smoked (per day) before becoming pregnant.
CIGARETTE_FIRST	0..1	float	A number of cigarettes the parent smoked (per day) during the first trimester.
CIGARETTE_SECOND	0..1	float	A number of cigarettes the parent smoked (per day) during the second trimester.
CIGARETTE_LAST	0..1	float	A number of cigarettes the parent smoked (per day) during the last trimester.
CIGARETTE_POST	0..1	float	A number of cigarettes the parent smoked (per day) postpartum.
DRINKS_PRE	0..1	float	A number of alcoholic drinks the parent consumed (per day) before becoming pregnant.
DRINKS_LAST	0..1	float	A number of alcoholic drinks the parent consumed (per day) during the last trimester.
GRAVIDA	0..1	Integer	A number of times the parent has been pregnant, including this pregnancy.
PARA	0..1	Integer	A number of viable pregnancies that had multiple fetuses.
PRE_PREGNANCY_WT	0..1	float	A measure of the parent's weight (in pounds) before becoming pregnant.
PRE_PREGNANCY_BMI	0..1	float	A measure of the parent's body mass index before becoming pregnant.
DELIVERY_WT	0..1	float	A measure of the parent's weight (in pounds) at delivery.
DELIVERY_PROCEDURE_ID	0..1	FK::PROCEDURES	A link to the procedures table (when applicable) that corresponds to the delivery.

Attribute	Cardinality	Type	Documentation
Outcome	0..*	FK::PREGNANCY_OUTCOME	A set of outcomes for this pregnancy.

PREGNANCY_OUTCOME

The PREGNANCY_OUTCOME table contains one record for each fetus resulting from the pregnancy.

Table 29. PREGNANCY_OUTCOME Details

Attribute	Cardinality	Type	Documentation
PREGNANCY_OUTCOME_ID	1	ID	
CHILDID	0..1	FK::DEMOGRAPHIC	A link back to the demographics table for the child (when applicable).
PARENTID	1	FK::PREGNANCY	A link to the information about the child's parent's pregnancy.
GESTATION_WEEKS	0..1	float	A number of weeks of gestation.
BIRTH_WT	0..1	float	A measure of the child's weight (in pounds) at birth.
BIRTH_HT	0..1	float	A measure of the child's length (in inches) at birth.
DISCHARGE_WT	0..1	float	A measure of the child's weight (in pounds) when discharged.
DISCHARGE_DATE	0..1	date	A date on which the child was discharged.
HAS_SMOKER_IN_HOUSE	0..1	Boolean	True if the child lives with an individual who smokes.
EVER_BREASTFED	0..1	Boolean	True if the child has ever been breastfed.
BREAST_FEEDING_FREQ	0..1	float	A number of times (per day) the child was breastfed, on average.
BREAST_FEEDING_STOPPED_AGE	0..1	float	An age of the child (in weeks) when breastfeeding stopped.
BREAST_FEEDING_STOPPED_REASON	0..1	String	A reason the child stopped breastfeeding. [TODO: Get the codes from WIC and decide if we're going to use those codes.]

PRESCRIBING

[From PCORnet CDM]

The PRESCRIBING table contains one record for each prescription ordered. Provider orders for medication dispensing and/or administration. These orders may take place in any setting, including the inpatient or outpatient basis.

Table 30. PRESCRIBING Details

Attribute	Cardinality	Type	Documentation
PRESCRIBINGID	1	ID	
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTER	Arbitrary encounter-level identifier. This should be present if the prescribing activity is directly associated with an encounter.
RX_PROVIDERID	0..1	FK::PROVIDER	Provider code for the provider who prescribed the medication. The provider code is a pseudoidentifier with a consistent crosswalk to the real identifier.
RX_ORDER_DATE	0..1	date	Order date of the prescription by the provider.
RX_ORDER_TIME	0..1	time	Order time of the prescription by the provider.
RX_START_DATE	0..1	date	Start date of order. This attribute may not be consistent with the date on which the patient actually begin taking the medication.
RX_END_DATE	0..1	date	End date of order (if available).
RX_DOSE_ORDERED	0..1	NUMERIC(x)	Dose of a given mediation, as ordered by the provider
RX_DOSE_ORDERED_UNIT	0..1	UNIT_TYPE	Units of measure associated with the dose of the medication as ordered by the provider
RX_QUANTITY	0..1	NUMERIC(x)	Quantity ordered.
RXNORM_CUI	0..1	CHAR (8)	Where an RxNorm mapping exists for the source medication, this field contains the RxNorm concept identifier (CUI) at the highest possible specificity.
RX_SOURCE	0..1	RX_ORIGIN_TYPE	Source of the prescribing information.

PROCEDURES

[From PCORnet CDM]

The PROCEDURES table contains one record per procedure for a patient. Procedure codes indicate the discrete medical interventions and diagnostic testing, such as surgical procedures and lab orders, delivered within a healthcare context.

Table 31. PROCEDURES Details

Attribute	Cardinality	Type	Documentation
PROCEDURESID	1	ID	
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTER	Arbitrary encounter-level identifier. Used to link across tables.
ENC_TYPE	0..1	ENCOUNTER_TYPE	This is a field replicated from the ENCOUNTER table. See the ENCOUNTER table for definitions.
ADMIT_DATE	0..1	date	This is a field replicated from the ENCOUNTER table. See the ENCOUNTER table for definitions.
PROVIDERID	0..1	FK::PROVIDER	Identifier of the PROVIDER most associated with the procedure order.
PX_DATE	0..1	date	Date the procedure was performed.
PX	1	CHAR (11)	Procedure code.
PX_TYPE	1	PX_TYPE	Procedure code type. We include a number of code types for flexibility, but the basic requirement that the code refer to a medical procedure remains.
PX_SOURCE	0..1	ORIGIN_TYPE	Source of the procedure information.
Immunization	0..*	FK::IMMUNIZATION	A set of immunizations associated with this procedure.
Pregnancy	0..1	FK::PREGNANCY	A pregnancy associated with a delivery procedure.

PROGRAM

The PROGRAM table contains one record for each distinct program. A program comprises a collection of interventions intended to produce a particular outcome.

Table 32. PROGRAM Details

Attribute	Cardinality	Type	Documentation
PROGRAMID	1	ID	
PROGRAM_NAME	1	String	A name of the program (e.g., Girls on the Run).
PROGRAM_DESCRIPTION	0..1	String	A description of the program.
PROGRAM_SETTING	0..1	SETTING_TYPE	A setting in which the program is offered (clinical or community).
AFFILIATED_PROGRAMID	0..1	FK::PROGRAM	A parent program which encompasses this and other programs that are often prescribed together. A parent program may not have any sessions of its own, nor a physical location but may serve only to encompass affiliated programs.
AIM_NUTRITION	0..1	Boolean	True if the aim of the program includes improving nutrition.
AIM_ACTIVITY	0..1	Boolean	True if the aim of the program includes improving physical activity.
AIM_WEIGHT	0..1	Boolean	True if the aim of the program includes improving weight status.
PRESCRIBED_TOTAL_DOSE	0..1	float	A total amount of time (in hours) an individual should spend in the program. This field should equal DURATION x FREQUENCY x LENGTH (weeks x sessions/week x hours/session).
PRESCRIBED_PROGRAM_DURATION	0..1	float	A measure of the time (in weeks) from start to finish.
PRESCRIBED_SESSION_FREQUENCY	0..1	float	A number of sessions delivered each week.
PRESCRIBED_SESSION_LENGTH	0..1	float	A number of hours delivered each session.
LOCATION_ADDRESS	0..1	String	A primary location at which this program's sessions are administered, expressed as an address. If the program is designed for participants to participate entirely on-line or at home, then the location value should be the string 'Virtual'.
LOCATION_LATITUDE	0..1	NUMERIC(8,6)	A latitude of the corresponding address location.
LOCATION_LONGITUDE	0..1	NUMERIC(9,6)	A latitude of the corresponding address location.
LOCATION_GEOCODE	0..1	varchar (15)	A primary location at which this program's sessions are administered, expressed as a geocode.
LOCATION_BOUNDARY_YEAR	0..1	NUMERIC(x)	A census year for which the corresponding geocode location applies.

Attribute	Cardinality	Type	Documentation
LOCATION_GEOLEVEL	0..1	GEOLEVEL_TYPE	A specificity of the geocode location. This can be assessed using logic that considers the length of the GEOCODE value (2 characters for state; 5 characters for county; 11 characters for census tract).
SESSION_OMISSION_PERCENT	0..1	float	A numeric estimate of the percentage of all sessions missing from the SESSION table (based on intended dose) for this program; 0% indicates a belief that the session information is fully populated.
SESSION_OMISSION_DESCRIPTION	0..1	String	A description of the circumstances under which session information for this program is missing; this field is required when the omission percent is greater than 0%.
SESSION_OMISSION_SYSTEMATIC	0..1	Boolean	True if session information for this program is systematically missing (e.g., because only half of the sessions are documented in an EHR).
AssetDelivery	0..*	FK::ASSET_DELIVERY	A set of asset deliveries associated with this program.
CurriculumComponent	0..*	FK::CURRICULUM_COMPONENT	A set of curriculum components for this program.
ProgramEnrollment	0..*	FK::PROGRAM_ENROLLMENT	A set of enrollments associated with this program.
Session	0..*	FK::SESSION	A set of sessions associated with this program.
ProgramReferral	0..*	FK::PROGRAM	A set of referrals to this program.

PROGRAM_ENROLLMENT

The PROGRAM_ENROLLMENT table contains one record for each person who enrolls in a program.

Table 33. PROGRAM_ENROLLMENT Details

Attribute	Cardinality	Type	Documentation
PROGRAM_ENROLLMENT_ID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographics table.

Attribute	Cardinality	Type	Documentation
PROGRAMID	1	FK::PROGRAM	A link back to the program this enrollment belongs to.
ENROLLMENT_DATE	0..1	date	A date on which the enrollment was performed.
COMPLETION_DATE	0..1	date	A date on which the individual who enrolled completed the program.
DISPOSITION_DESCRIPTION	0..1	String	A description of the circumstances under which an individual ended their participation in the program. For example, an individual might complete a program successfully, they might drop out, or they might move to a different state.

PROVIDER

[From PCORnet CDM]

The PROVIDER table contains one record per PROVIDER ID. Data about the providers who are involved in the care processes documented in the CDM.

Table 34. PROVIDER Details

Attribute	Cardinality	Type	Documentation
PROVIDERID	1	ID	
PROVIDER_SEX	0..1	SEX_TYPE	Sex assigned at birth.
PROVIDER_SPECIALTY_PRIMARY	0..1	SPECIALTY_TYPE	Primary specialty of the provider
Diagnosis	0..*	FK::DIAGNOSIS	A set of diagnoses associated with this provider.
Encounter	0..*	FK::ENCOUNTER	A set of encounters associated with this provider.
Immunization	0..*	FK::IMMUNIZATION	A set of immunizations associated with this provider.
Prescribing	0..*	FK::PRESCRIBING	A set of prescriptions ordered by this provider.
Procedure	0..*	FK::PROCEDURES	A set of procedures ordered by this provider.
Referral	0..*	FK::REFERRAL	A set of referrals made by this provider.

Attribute	Cardinality	Type	Documentation
Session	0..*	FK::SESSION	A set of sessions associated with this provider.

PRO_CM

[From PCORnet CDM]

This table is used to store responses to patient-reported outcome measures (PROs) or questionnaires. This table can be used to store item-level responses as well as the overall score for each measure.

Table 35. PRO_CM Details

Attribute	Cardinality	Type	Documentation
PRO_CM_ID	1	ID	
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier for the patient for whom the PRO response was captured. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTER	Arbitrary encounter-level identifier used to link across tables. This is an optional field, and should only be populated if the item was collected as part of a healthcare encounter.
PRO_DATE	1	date	The date of the response.
PRO_TIME	0..1	time	The time of the response.
PRO_TYPE	0..1	PRO_TYPE_TYPE	Terminology / vocabulary used to describe the PRO item. More information on PROMIS, Neuro-QoL and ASQC-Me and the NIH Toolbox can be found on the HealthMeasures website. (www.healthmeasures.net) The Patient-Reported Outcome version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™) is maintained by the National Cancer Institute. (https://healthcaredelivery.cancer.gov/pro-ctcae/) Information on the Hospital Consumer Assessment of Healthcare Providers and Systems (HPCAHPS) is located here: http://www.hcahpsonline.org
PRO_ITEM_NAME	0..1	String	Short name or code of the PRO item in the vocabulary/terminology specified in PRO_TYPE.
PRO_ITEM_LOINC	0..1	VARCHAR (10)	LOINC® code for the PRO item, if available.

Attribute	Cardinality	Type	Documentation
			Logical Observation Identifiers, Names, and Codes (LOINC) from the Regenstrief Institute. Current LOINC codes are from 3-7 characters long but Regenstrief suggests a length of 10 for future growth. The last digit of the LOINC code is a check digit and is always preceded by a hyphen. All parts of the LOINC code, including the hyphen, must be included. Do not pad the LOINC code with leading zeros.
PRO_RESPONSE_TEXT	0..1	String	Text version of the response recorded for the item, if available/applicable.
PRO_RESPONSE_NUM	0..1	NUMERIC(x)	The numeric response recorded for the item, if available/applicable.
PRO_METHOD	0..1	PRO_METHOD_TYPE	Method of administration. Electronic includes responses captured via a personal or tablet computer, at web kiosks, or via a smartphone.
PRO_MODE	0..1	PRO_MODE_TYPE	The person who responded on behalf of the patient for whom the response was captured. A proxy report is a measurement based on a report by someone other than the patient reporting as if he or she is the patient, such as a parent responding for a child, or a caregiver responding for an individual unable to report for themselves. Assistance excludes providing interpretation of the patient's response.
PRO_CAT	0..1	PRO_CAT_TYPE	Indicates whether Computer Adaptive Testing (CAT) was used to administer the survey or instrument that the item was part of. May apply to electronic (EC) and telephonic (PH or IV) modes.
PRO_SOURCE	0..1	PRO_SOURCE_TYPE	Source of the information for the PRO result.
PRO_ITEM_VERSION	0..1	String	Version of the item/question.
PRO_MEASURE_NAME	0..1	String	Short name or code of the PRO measure/form that item belongs to, if item is being administered as part of a measure
PRO_MEASURE_SEQ	0..1	String	Arbitrary ID/sequence number used to link PRO item responses that are associated with the same measure/form.
PRO_MEASURE_SCORE	0..1	NUMERIC(x)	Overall raw score for the PRO measure.
PRO_MEASURE_THETA	0..1	NUMERIC(x)	The value of theta reported from the CAT PROMIS results. Only applies to items that are administered as part of a measure.
PRO_MEASURE_SCALED_TSCORE	0..1	NUMERIC(x)	Standardized score based on the total raw score for the instrument. Only applies to items that are administered as part of a measure.

Attribute	Cardinality	Type	Documentation
PRO_MEASURE_STANDARD_ERROR	0..1	NUMERIC(x)	Possible range of the actual final score based on the scaled T-score. Only applies to items that are administered as part of a measure.
PRO_MEASURE_COUNT_SCORED	0..1	NUMERIC(x)	Number of PRO item responses that were involved in the scoring of the measure.
PRO_MEASURE_LOINC	0..1	VARCHAR (10)	LOINC® code for the PRO item, if available. Logical Observation Identifiers, Names, and Codes (LOINC) from the Regenstrief Institute. Current LOINC codes are from 3-7 characters long but Regenstrief suggests a length of 10 for future growth. The last digit of the LOINC code is a check digit and is always preceded by a hyphen. All parts of the LOINC code, including the hyphen, must be included. Do not pad the LOINC code with leading zeros.
PRO_MEASURE_VERSION	0..1	String	Version of the measure.
PRO_ITEM_FULLNAME	0..1	String	Full name of the PRO item.
PRO_ITEM_TEXT	0..1	String	Text of the PRO item question.
PRO_MEASURE_FULLNAME	0..1	String	Full name of the PRO measure. (CODI note: Used for “raw” or original source name for nonstandard measures.)

REFERRAL

The REFERRAL table contains one record for each outgoing or incoming referral to a clinical provider or a program.

Table 36. REFERRAL Details

Attribute	Cardinality	Type	Documentation
REFERRALID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographic table.
ENCOUNTERID	0..1	FK::ENCOUNTER	A link back to the encounter table, if the referral can be unambiguously associated with an encounter.
DIRECTION	1	DIRECTION_TYPE	An indication of whether the referral was incoming or outgoing.
REFERRAL_DATE	1	date	A date the referral was made.
REFERRAL_STATUS	0..1	REFERRAL_STATUS_TYPE	A final disposition of the referral.

Attribute	Cardinality	Type	Documentation
REFERRAL_PRIOR_AUTH	0..1	YES//NO	An indication of whether prior authorization was required for the referral.
SOURCE_PROVIDER_ID	0..1	FK::PROVIDER	A provider responsible for initiating this referral.
SOURCE_ORGANIZATION	0..1	ORGANIZATION_TYPE	An organization that initiated the referral.
DESTINATION_ORGANIZATION	0..1	ORGANIZATION_TYPE	An organization to which the referral was sent.
DESTINATION_SPECIALTY	0..1	SPECIALTY_TYPE	A clinical specialty for which the patient is being referred.
DESTINATION_PROGRAMID	0..1	FK:PROGRAM	A program to which an individual is being referred.

SDOH_EVIDENCE_INDICATOR

The SDOH_EVIDENCE_INDICATOR table contains one record for each piece of SDOH relevant data. It conveys the existence of information regarding social circumstance(s) considered to be a determinant of health for an individual patient or program participant. The information may come from an administered SDOH screening or diagnosis, or problem associated with the individual and is relevant to one of the SDOH categories (e.g., FOOD_DOMAIN, FINANCIAL_DOMAIN). The purpose of the evidence indicator is to provide a short-cut for knowing if relevant information exists and does not indicate whether a social risk exists.

Table 37. SDOH_EVIDENCE_INDICATOR Details

Attribute	Cardinality	Type	Documentation
SDOH_EVIDENCE_INDICATOR_ID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographics table.
SDOH_CATEGORY	1	SDOH_CATEGORY_TYPE	A social topic area pertaining to circumstances which can determine health outcomes for an individual.
EVIDENCE_DATE	1	date	A date on which a data owner, partner, or researcher has made an assertion indicating the presence of SDOH evidence. This date corresponds to the data partner's most recent determination of available evidence and does not necessarily match submission dates of any of the SDOH evidence. CODI is not expected to maintain a history of assertions, only one assertion based on the data partner's supplied evidence.

Attribute	Cardinality	Type	Documentation
EVIDENCE_TABLE_NAME	0..1	String	A name of a table in the CODI schema in which there is some evidence pertaining to the CODI SDOH indicator category. The evidence may be a screening response (in PRO_CM), or a reported problem (in CONDITION or DIAGNOSIS), or some other information stored in a CODI table.
EVIDENCE_EXPLANATION	0..1	String	For indicator assertions without CODI data evidence; an explanation for the assertion.
EVIDENCE_ROWID	0..1	ID	An identifier for a specific row in the table referenced in the EVIDENCE_TABLE_NAME that contains evidence of a potential social determinant.

SESSION

The SESSION table contains one record for each session. A session is a specific point in time where an individual or family is involved in programming that focuses on the prevention or intervention of chronic disease, or chronic-related comorbidities.

In a clinical setting, a session corresponds to a visit. There may be multiple visits in a single encounter. The ENCOUNTERID field is required for clinical sessions.

In a community setting, a session corresponds to one component of a program. The PROGRAMID field is required for sessions that are components of a program.

At least one of those fields should be present in every case.

Table 38. SESSION Details

Attribute	Cardinality	Type	Documentation
SESSIONID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographics table.
ENCOUNTERID	0..1	FK::ENCOUNTER	A link back to the encounter this session corresponds to (if any).
PROVIDERID	0..1	FK::PROVIDER	A provider primarily responsible for this session.
PROGRAMID	0..1	FK::PROGRAM	A link back to the program this session belongs to (if any).
SESSION_DATE	0..1	date	A date on which the session was conducted.

Attribute	Cardinality	Type	Documentation
SESSION_MODE	0..1	MODE_TYPE	An indication of the way the session was delivered (e.g., individual, group, phone).
SCREENING	0..1	PROCESS_PERFORMED_TYPE	True if the session included any assessment of lifestyle behaviors related to obesity, such as physical activity, nutrition, screen time, or sleep.
COUNSELING	0..1	PROCESS_PERFORMED_TYPE	True if the session included any advice or direction regarding lifestyle related to obesity, such as physical activity, nutrition, screen time, or sleep.
INTERVENTION_ACTIVITY	0..1	PROCESS_PERFORMED_TYPE	True if the session included performing at least moderate physical activity; moderate activity requires a moderate amount of effort (5-6 on a scale of 0 to 10) and noticeably accelerates the heart rate and breathing.
INTERVENTION_NUTRITION	0..1	PROCESS_PERFORMED_TYPE	True if the session included an activity designed to improve nutrition.
INTERVENTION_NAVIGATION	0..1	PROCESS_PERFORMED_TYPE	True if the session included a navigational service to access benefits or to overcome barriers to care.
DOSE	0..1	float	A measure of the amount of time spent on this encounter. Researchers can compare the total dose to the prescribed total dose to assess the extent to which an individual completed a program.
CURRICULUM_COMPONENT_ID	0..1	FK::CURRICULUM_COMPONENT	A link back to the curriculum component (if any) associated with this session.
Alert	0..*	FK::SESSION_ALERT	A set of triggered alerts associated with this session.

SESSION_ALERT

The SESSION_ALERT table contains one record for each alert that triggered during a session.

Table 39. SESSION_ALERT Details

Attribute	Cardinality	Type	Documentation
SESSION_ALERT_ID	1	ID	
ALERTID	1	FK::ALERT	An alert that triggered during a session.
SESSIONID	1	FK::SESSION	A session during which an alert triggered.

Attribute	Cardinality	Type	Documentation
ALERT_DATE	0..1	date	A date that an alert triggered.
ALERT_TIME	0..1	time	A time that an alert triggered.

VITAL

[From PCORnet CDM]

The VITAL table contains one record for each measurement of vital signs. Vital signs (such as height, weight, and blood pressure) directly measure an individual's current state of attributes.

Table 40. VITAL Details

Attribute	Cardinality	Type	Documentation
VITALID	1	ID	
PATID	1	FK::DEMOGRAPHIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTER	Arbitrary encounter-level identifier. Not all vital sign measures will be associated with a healthcare encounter.
MEASURE_DATE	1	date	Date of vitals measure.
MEASURE_TIME	0..1	time	Time of vitals measure.
VITAL_SOURCE	1	VITAL_SOURCE_TYPE	Please note: The "Patient-reported" category can include reporting by patient's family or guardian.
HT	0..1	NUMERIC(x)	Height (in inches) measured by standing. Only populated if measure was taken on this date. If missing, this value should be null. Decimal precision is permissible.
WT	0..1	NUMERIC(x)	Weight (in pounds). Only populated if measure was taken on this date. If missing, this value should be null. Decimal precision is permissible.
DIASTOLIC	0..1	NUMERIC(x)	Diastolic blood pressure (in mmHg). Only populated if measure was taken on this date. If missing, this value should be null.
SYSTOLIC	0..1	NUMERIC(x)	Systolic blood pressure (in mmHg). Only populated if measure was taken on this date. If missing, this value should be null.
ORIGINAL_BMI	0..1	NUMERIC(x)	BMI if calculated in the source system. Decimal precision is permissible.

Appendix E CODI Research Data Model Codeset Details

The following tables contain the codeset values (aka value sets) and meanings for all the fields whose values are constrained by a finite set of code values in the CODI data models. Most of these code sets are adopted from PCORnet Common Data Model Version 6.0. Any code sets that are adopted are marked as such. This appendix is not the source of record for adopted code sets. They are listed here for convenience.

Implementers should be aware of changes to set values in adopted models by referring to the adopted model's official data model. The impact to CODI and CODI ancillary tables from updates to the adopted models must be evaluated and may require an update to CODI's implementation guide.

ABN_TYPE

[From PCORnet CDM]

Table 41. ABN_TYPE Details

Code	Documentation
AB	Abnormal
AH	Abnormally high
AL	Abnormally low
CH	Critically high
CL	Critically low
CR	Critical
IN	Inconclusive
NL	Normal
NI	No information
UN	Unknown
OT	Other

ADDRESS_TYPE_TYPE

[From PCORnet CDM]

Table 42. ADDRESS_TYPE_TYPE Details

Code	Documentation
PO	Postal
PH	Physical
NI	No information
UN	Unknown
OT	Other

ADDRESS_USE_TYPE

[From PCORnet CDM]

Table 43: ADDRESS_USE_TYPE Details

Code	Documentation
HO	Home
WO	Work
TP	Temp
OL	Old/Incorrect
NI	No information
UN	Unknown
OT	Other

ADMITTING_SOURCE_TYPE

[From PCORnet CDM]

Table 44. ADMITTING_SOURCE_TYPE Details

Code	Documentation
AF	Adult Foster Home
AL	Assisted Living Facility
AV	Ambulatory Visit
ED	Emergency Department
HH	Home Health
HO	Home / Self Care
HS	Hospice
IP	Other Acute Inpatient Hospital
NH	Nursing Home (Includes ICF)
RH	Rehabilitation Facility
RS	Residential Facility
SN	Skilled Nursing Facility
IH	Intra-hospital
NI	No information
UN	Unknown
OT	Other

ASSET_TYPE

This codeset enumerates different kinds of assets a person might receive.

Table 45. ASSET_TYPE Details

Code	Documentation
CC	Childcare
FO	Food
HI	Health insurance
TR	Transportation
PH	Permanent Housing
NI	No information
UN	Unknown
OT	Other

CONDITION_SOURCE_TYPE

[From PCORnet CDM]

Table 46. CONDITION_SOURCE_TYPE Details

Code	Documentation
PR	Patient-reported medical history
HC	Healthcare problem list
RG	Registry cohort
CC	Patient Chief Complaint
PC	PCORnet-defined condition algorithm
DR	Derived
NI	No information
UN	Unknown
OT	Other

CONDITION_STATUS_TYPE

[From PCORnet CDM]

Table 47. CONDITION_STATUS_TYPE Details

Code	Documentation
AC	Active
RS	Resolved
IN	Inactive
NI	No information
UN	Unknown
OT	Other

CONDITION_TYPE_TYPE

[From PCORnet CDM]

Table 48. CONDITION_TYPE_TYPE Details

Code	Documentation
09	ICD-9-CM
10	ICD-10-CM
11	ICD-11-CM
SM	SNOMED CT
HP	Human Phenotype Ontology
AG	Algorithmic
NI	No information
UN	Unknown
OT	Other

DATAMART_EHR_TYPE

[From PCORnet CDM]

Table 49. DATAMART_EHR_TYPE Details

Attribute	Documentation
01	Not present
02	Present
NI	No information
UN	Unknown
OT	Other

DATAMART_PLATFORM_TYPE

[From PCORnet CDM]

Table 50. DATAMART_PLATFORM_TYPE Details

Attribute	Documentation
01	SQL Server
02	Oracle
03	PostgreSQL
04	MySQL
05	SAS
NI	No information
UN	Unknown
OT	Other

DIRECTION_TYPE

This codeset enumerates the direction of a referral.

Table 51. DIRECTION_TYPE Details

Code	Documentation
I	Incoming: A referral to this data contributor.
O	Outgoing: A referral from this data contributor.

DISCHARGE_STATUS_TYPE

[From PCORnet CDM]

Table 52. DISCHARGE_STATUS_TYPE Details

Code	Documentation
AF	Adult Foster Home
AL	Assisted Living Facility
AM	Against Medical Advice
AW	Absent without leave
EX	Expired
HH	Home Health
HO	Home / Self Care
HS	Hospice
IP	Other Acute Inpatient Hospital
NH	Nursing Home (Includes ICF)
RH	Rehabilitation Facility
RS	Residential Facility
SH	Still In Hospital
SN	Skilled Nursing Facility
NI	No information
UN	Unknown
OT	Other

DISCHARGE_TYPE

[From PCORnet CDM]

Table 53. DISCHARGE_TYPE Details

Code	Documentation
A	Discharged alive
E	Expired
NI	No information
UN	Unknown
OT	Other

DRG_TYPE

[From PCORnet CDM]

Table 54. DRG_TYPE Details

Code	Documentation
01	CMS-DRG (old system)
02	MS-DRG (current system)
NI	No information
UN	Unknown
OT	Other

DX_SOURCE_TYPE

[From PCORnet CDM]

Table 55. DX_SOURCE_TYPE Details

Code	Documentation
AD	Admitting
DI	Discharge
FI	Final
IN	Interim
NI	No information
UN	Unknown
OT	Other

DX_TYPE

[From PCORnet CDM]

Table 56. DX_TYPE Details

Code	Documentation
09	ICD-9-CM
10	ICD-10-CM
11	ICD-11-CM
SM	SNOMED CT
NI	No information
UN	Unknown
OT	Other

ENCOUNTER_TYPE

[From PCORnet CDM]

Table 57. ENCOUNTER_TYPE Details

Code	Documentation
AV	Ambulatory Visit: Includes visits at outpatient clinics, physician offices, same day/ambulatory surgery centers, urgent care facilities, and other same-day ambulatory hospital encounters, but excludes emergency department encounters.
ED	Emergency Department (ED): Includes ED encounters that become inpatient stays (in which case inpatient stays would be a separate encounter). Excludes urgent care facility visits. ED claims should be pulled before hospitalization claims to ensure that ED with subsequent admission won't be rolled up in the hospital event. Does not include observation stays, where known.
EI	Emergency Department Admit to Inpatient Hospital Stay: Permissible substitution for preferred state of separate ED and IP records. Only for use with data sources where the individual records for ED and IP cannot be distinguished.
IP	Inpatient Hospital Stay: Includes all inpatient stays, including: same-day hospital discharges, hospital transfers, and acute hospital care where the discharge is after the admission date. Does not include observation stays, where known.
IS	Non-Acute Institutional Stay: Includes hospice, skilled nursing facility (SNF), rehab center, nursing home, residential, overnight non-hospital dialysis, and other non-hospital stays.
OS	Observation Stay: “Hospital outpatient services given to help the doctor decide if the patient needs to be admitted as an inpatient or can be discharged. Observations services may be given in the emergency department or another area of the hospital.” Definition from Medicare, CMS Product No. 11435, https://www.medicare.gov/Pubs/pdf/11435.pdf.
IC	Institutional Professional Consult: Permissible substitution when services provided by a medical professional cannot be combined with the given encounter record, such as a specialist consult in an inpatient setting; this situation can be common with claims data sources. This includes physician consults for patients during inpatient encounters that are not directly related to the cause of the admission (e.g. a ophthalmologist consult for a patient with diabetic ketoacidosis) guidance updated in v4.0).
OA	Other Ambulatory Visit: Includes other non-overnight AV encounters such as hospice visits, home health visits, skilled nursing visits, other non-hospital visits, as well as telemedicine, telephone and email consultations. May also include "lab only" visits (when a lab is ordered outside of a patient visit), "pharmacy only" (e.g., when a patient has a refill ordered without a face-to-face visit), "imaging only", etc.
NI	No information
UN	Unknown
OT	Other

FACILITY_TYPE

[From PCORnet CDM]

See CDM Value Set Appendix for a list of acceptable values.

Table 58. FACILITY_TYPE Details

Code	Documentation
Refer to the value set appendix to the PCORnet CDM documentation	

FREQ_TYPE

This codeset enumerates different frequencies that something happens.

Table 59. FREQ_TYPE Details

Code	Documentation
O	Once
D	Daily
W	Weekly
M	Monthly
Y	Yearly

GENDER_IDENTITY_TYPE

[From PCORnet CDM]

Table 60. GENDER_IDENTITY_TYPE Details

Code	Documentation
M	Man
F	Woman
TM	Transgender male/Trans man/Female-to-male
TF	Transgender female/Trans woman/Male-to-female
GQ	Genderqueer/Non-Binary
SE	Something else
MU	Multiple gender categories
DC	Decline to answer
NI	No information
UN	Unknown
OT	Other

GEOLEVEL_TYPE

[From CHORDS VDW]

Table 61. GEOLEVEL_TYPE Details

Code	Documentation
B	Block

Code	Documentation
G	Block Group
T	Census Track
C	County
Z	Zip Code
P	Post Office
U	Unknown, unable to append

LANGUAGE_TYPE

[From PCORnet CDM]

See CDM Value Set Appendix for a list of acceptable values.

Table 62. LANGUAGE_TYPE Details

Code	Documentation
Refer to the value set appendix in the PCORnet CDM documentation	

MODE_TYPE

This codeset enumerates different ways in which interventions can be delivered.

Table 63. MODE_TYPE Details

Code	Documentation
I	Individual Meeting
G	Group Meeting
W	Web
T	Telephone
M	Mail

MODIFIER_TYPE

[From PCORnet CDM]

Table 64. MODIFIER_TYPE Details

Code	Documentation
EQ	Equal
GE	Greater than or equal to
GT	Greater than
LE	Less than or equal to
LT	Less than
TX	Text

Code	Documentation
NI	No information
UN	Unknown
OT	Other

ORGANIZATION_TYPE

For clinical organizations, use the CMS Certification Number (CCN); each implementing network will need to choose a representative CCN for its clinical data partners. For community organizations, each implementing network will need to establish a set of community organization codes. These additional codes should include at least one letter so that they do not conflict with CCNs.

Table 65. ORGANIZATION_TYPE Details

Code	Documentation
Refer to the CMS Certification Number note above.	

ORIGIN_TYPE

[From PCORnet CDM]

Table 66. ORIGIN_TYPE Details

Code	Documentation
OD	Order/EHR
BI	Billing
CL	Claim
DR	Derived
NI	No information
UN	Unknown
OT	Other

PAYER_TYPE

[From PCORnet CDM

See CDM Value Set Appendix for a list of acceptable values.

Table 67. PAYER_TYPE Details

Code	Documentation
Refer to the value set appendix to the PCORnet CDM documentation	

PROCESS_PERFORMED_TYPE

This codeset enumerates the extent to which process steps might be conducted. Process steps include screening, counseling, and interventions

Table 68. PROCESS_PERFORMED_TYPE Details

Code	Documentation
Y	Yes, the process step was conducted.
N	No, the process step was not conducted.
NI	No information
UN	Unknown
OT	Other

PRO_CAT_TYPE

[From PCORnet CDM]

Table 69. PRO_CAT_TYPE Details

Code	Documentation
Y	Yes
N	No
NI	No Information
UN	Unknown
OT	Other

PRO_METHOD_TYPE

[From PCORnet CDM]

Table 70. PRO_METHOD_TYPE Details

Code	Documentation
PA	Paper
EC	Electronic
PH	Telephonic
IV	Telephonic with interactive voice response (IVR) technology
NI	No information

Code	Documentation
UN	Unknown
OT	Other

PRO_MODE_TYPE

[From PCORnet CDM]

Table 71. PRO_MODE_TYPE Details

Code	Documentation
SF	Self without assistance
SA	Self with assistance
PR	Proxy without assistance
PA	Proxy with assistance
NI	No information
UN	Unknown
OT	Other

PRO_SOURCE_TYPE

[From PCORnet CDM]

Table 72. PRO_SOURCE_TYPE Details

Code	Documentation
OD	Order/EHR
BI	Billing
CL	Claim
SR	Survey system/mobile app
DR	Derived
NI	No information
UN	Unknown
OT	Other

PRO_TYPE_TYPE

[From PCORnet CDM]

Table 73. PRO_TYPE_TYPE Details

Code	Documentation
NQ	Neuro-QoL
PM	PROMIS
AM	ASQC-Me

Code	Documentation
NT	NIH Toolbox
PC	PRO_CTCAH
LC	LOINC
HC	HCAHPS
NI	No Information
UN	Unknown
OT	Other

PX_TYPE

[From PCORnet CDM]

Table 74. PX_TYPE Details

Code	Documentation
09	ICD-9-CM
10	ICD-10-PCS
11	ICD-11-PCS
CH	CPT or HCPCS
LC	LOINC
ND	NDC
RE	Revenue
NI	No information
UN	Unknown
OT	Other

QUAL_TYPE

[From PCORnet CDM]

See CDM Value Set Appendix for a list of acceptable values.

Table 75. QUAL_TYPE Details

Code	Documentation
Refer to the value set appendix to the PCORnet CDM documentation	

RACE_TYPE

[From PCORnet CDM]

Table 76. RACE_TYPE Details

Code	Documentation
01	American Indian or Alaska Native
02	Asian
03	Black or African American
04	Native Hawaiian or Other Pacific Islander
05	White
06	Multiple race
07	Refuse to answer
NI	No information
UN	Unknown
OT	Other

REFERRAL_STATUS_TYPE

[From PCORnet CDM]

Table 77. REFERRAL_STATUS_TYPE Details

Code	Documentation
A	Approved
D	Denied
NI	No information
UN	Unknown
OT	Other

RELATIONSHIP_TYPE

See https://www.hl7.org/fhir/valueset-relatedperson-relationshiptype.html

Table 78. RELATIONSHIP_TYPE Details

Code	Documentation
Refer to https://www.hl7.org/fhir/valueset-relatedperson-relationshiptype.html	

RESULT_LOC_TYPE

[From PCORnet CDM]

Table 79. RESULT_LOC_TYPE Details

Code	Documentation
L	Lab
P	Point of Care

Code	Documentation
NI	No information
UN	Unknown
OT	Other

RX_ORIGIN_TYPE

[From PCORnet CDM]

Table 80. RX_ORIGIN_TYPE Details

Code	Documentation
OD	Order/EHR
DR	Derived
NI	No information
UN	Unknown
OT	Other

SDOH_CATEGORY_TYPE

This codeset enumerates categories of social factors related to the same topic (e.g., housing, food) that can determine health outcomes. These align with the harmonization effort by the Gravity Project.

Table 81. SDOH_CATEGORY_TYPE Details

Code	Documentation
FD	FOOD_DOMAIN - Pertaining to an individual's access to adequate, nutritional, safe, and culturally acceptable food.
HS	HOUSING_STABILITY_DOMAIN - Pertaining to an individual's access to temporary or permanent reliable shelter.
HA	HOUSING_ADEQUACY_DOMAIN - Pertaining to the habitability of an individual's housing.
TR	TRANSPORTATION_DOMAIN - Pertaining to an individual's access to transportation for routine life sustaining activities such as to place of employment, medical facilities, and school.
IV	INTERPERSONAL_VIOLENCE_DOMAIN - Pertaining to an individual's physical and emotional safety in close relationships.
FI	FINANCIAL_DOMAIN - Pertaining to an individual's ability to or feeling about meeting current and/or ongoing financial obligations.
MN	MATERIAL_NECESSESITIES_DOMAIN - Pertaining to an individual's access to socially perceived physical necessities.
EM	EMPLOYMENT_DOMAIN - Pertaining to an individual's status on having, looking for, or being without a job or work.
HI	HEALTH_INSURANCE_DOMAIN - Pertaining to an individual's access to health insurance.

Code	Documentation
EC	ELDER_CARE_DOMAIN - Pertaining to an elder's exposure to physical, psychological, sexual, or financial abuse, or neglect by caregivers.
ED	EDUCATION_DOMAIN - Pertaining to an individual's academic achievements.
ST	STRESS_DOMAIN - Pertaining to an individual's ability to meet, mitigate, or alter perceived excesses in environmental demands and stimuli.
VE	VETERAN_DOMAIN - Pertaining to an individual's current and historical status in military service.
SC	SOCIAL_CONNECTION_DOMAIN - Pertaining to an individual's actual or perceived frequency of social contact, and actual or perceived access to informational, tangible, and emotional support from others.

SETTING_TYPE

This codeset enumerates different settings in which interventions can be delivered.

Table 82. SETTING_TYPE Details

Code	Documentation
CL	Clinical: Healthcare organization that provides clinical services to patients. [Pulled from glossary]
CO	Community: Setting in which weight-related services or assets are delivered that is NOT a clinic, hospital, health center, or other site of clinical care; examples include YMCA, Boys & Girls Club, Parks & Rec sites [Pulled from glossary]

SEX_TYPE

[From PCORnet CDM]

Table 83. SEX_TYPE Details

Code	Documentation
A	Ambiguous
F	Female
M	Male
NI	No information
UN	Unknown
OT	Other

SPECIALTY_TYPE

See http://nucc.org/index.php/code-sets-mainmenu-41/provider-taxonomy-mainmenu-40/pdf-mainmenu-53

Table 84. SPECIALTY_TYPE Details

Code	Documentation
Refer to http://nucc.org/index.php/code-sets-mainmenu-41/provider-taxonomy-mainmenu-40/pdf-mainmenu-53	

SPECIMEN_SOURCE_TYPE

[From PCORnet CDM]

See CDM Value Set Appendix for a list of acceptable values.

Table 85. SPECIMEN_SOURCE_TYPE Details

Code	Documentation
Refer to the value set appendix to the PCORnet CDM documentation	

STATE

A 2 character code for a state or territory of the United States of America. For example, AL for Alabama. See the CDM Value Set Appendix for a list of acceptable values.

Table 86. STATE Details

Code	Documentation
Refer to the value set appendix to the PCORnet CDM documentation	

UNIT_TYPE

[From PCORnet CDM]

See CDM Value Set Appendix for a list of acceptable values.

Table 87. UNIT_TYPE Details

Code	Documentation
Refer to the value set appendix to the PCORnet CDM documentation	

VITAL_SOURCE_TYPE

[From PCORnet CDM]

Table 88. VITAL_SOURCE_TYPE Details

Code	Documentation
PR	Patient-reported
PD	Patient device direct feed
HC	Healthcare delivery setting
HD	Healthcare device direct feed
DR	Derived
NI	No information
UN	Unknown
OT	Other

VX_CODE_TYPE_TYPE

[From PCORnet CDM]

Table 89. VX_CODE_TYPE_TYPE Details

Code	Documentation
CX	CVX
ND	NDC
CH	CPT or HCPCS
RX	RXNORM
NI	No information
UN	Unknown
OT	Other

VX_MANUFACTURER_TYPE

[From PCORnet CDM] See CDM Value Set documentation for manufacturer code list.

Table 90. VX_MANUFACTURER_TYPE Details

Code	Documentation
Refer to the value set appendix to the PCORnet CDM documentation	

VX_SOURCE_TYPE

[From PCORnet CDM]

Table 91. VX_SOURCE_TYPE Details

Code	Documentation
OD	International administration
EF	External Feed
IS	Immunization Information System
PR	Patient-reported
DR	Derived
NI	No information
UN	Unknown
OT	Other

VX_STATUS_REASON_TYPE

[From PCORnet CDM]

Table 92. VX_STATUS_REASON_TYPE Details

Code	Documentation
IM	Immunity
MP	Medical precaution
OS	Out of stock
PO	Patient objection
NI	No information
UN	Unknown
OT	Other

VX_STATUS_TYPE

[From PCORnet CDM]

Table 93. VX_STATUS_TYPE Details

Code	Documentation
CP	Completed
ER	Entered in error

Code	Documentation
ND	Not Done
IC	Incomplete
NI	No information
UN	Unknown
OT	Other

YES//NO

[From PCORnet CDM]

Table 94. YES//NO Details

Code	Documentation
Y	Yes
N	No
R	Refuse to answer
NI	No information
UN	Unknown
OT	Other

Acronyms

Term	Definition
ACS	American Community Survey
BMI	Body Mass Index
CAT	Computer Adaptive Testing
CBO	Community-Based Organization
CCLA	Clinical Community Linkages Assessment
CCN	CMS Certification Number
CCWG	CODI Collaborative Work Group
CDC	Centers for Disease Control and Prevention
CDM	Common Data Model
CDS	Clinical Decision Support System
CHORDS	Colorado Health Observation Regional Data Service
CMS	Centers for Medicare & Medicaid Services
CODI	Clinical and Community Health Data Initiative (formally Childhood Obesity Data Initiative)
DM IG	Data Models Implementation Guide
DCC	Data Coordinating Center
EHR	Electronic Health Record
ETL	Extract–Transform–Load
FFRDC	Federally Funded Research and Development Center
FHIR	Fast Health Information Resource
HL7	Health Level Seven International
HPCAHPS	Hospital Consumer Assessment of Healthcare Providers and Systems
ICD	International Clinical Diagnosis
IT	Information Technology
LOINC	Logical Observation Identifiers, Names, and Codes
OMOP	Observational Medical Outcomes Partnership
PCORnet	Patient Centered Outcomes Research Network
PII	Personally Identifiable Information
PPRL	Privacy-Preserving Record Linkage

PRO	Patient-Reported Outcome
PRO-CTCAE	Patient-Reported Outcome version of the Common Terminology Criteria for Adverse Events
RDM	Research Data Model
RLDM	Record Linkage Data Model
RUCA	Rural-Urban Commuting Area
SDOH	Social Determinants of Health
SNOMED	Systematized Nomenclature of Human Medicine
TES	Technical Environmental Scan
VDW	Virtual Data Warehouse
WIC	Women, Infants, and Children

Resources

Centers for Disease Control and Prevention. *Growth Chart Training*. Available: https://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm

The Colorado Health Observation Regional Data Service. *CHORDS VDW 3.5 Data Model Manual*, March 2021. Denver, CO. Available by sending a request to CODI@cdc.gov.

HL7. *HL7 v3 Code System NullFlavor.* Available: https://www.hl7.org/fhir/v3/NullFlavor/cs.html

HL7 & Boston Children's Hospital. *CDC Hooks*, 2018. Available: https://cds-hooks.org/

PCORnet. *PCORnet Common Data Model v6.0 Specification*, October 2020. Available: https://pcornet.org/wp-content/uploads/2020/12/PCORnet-Common-Data-Model-v60-2020_10_221.pdf

V. Rogers, P. Hart, E. Motyka, E. Rines, J. Vine, D. Deatrick. *Impact of Let's Go! 5-2-1-0: a Community-based, Multisetting Childhood Obesity Prevention Program*, 2013. J Pediatr Psychol (9): 1010–20.

C. Reich, P. Ryan, R. Belenkaya, K. Natarajan, C. Blacketer, *OMOP Common Data Model v6.0 Specifications*, Oct. 11, 2018. Available: https://github.com/OHDSI/CommonDataModel/wiki

Wikipedia. *Cryptographic hash function.* Available: https://en.wikipedia.org/wiki/Cryptographic_hash_function

Health Federally Funded Research and Development Center. *CODI Privacy Preserving Record Linkage Implementation Guide*, Version 2.1, September 1, 2022. Available: https://github.com/mitre/codi/blob/main/CODI%20PPRL%20Implementation%20Guide.pdf

Health Federally Funded Research and Development Center. *The Clinical and Community Data Initiative Gaps Analysis*, August 02, 2021. Available by contacting Division of Nutrition, Physical Activity, and Obesity at CDC.gov for access

NOTICE