

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**

Childhood Obesity Data Initiative

Task Order No. 200-2016-F-89363

CODI Data Models Implementation Guide

For the Denver, Colorado Pilot (2019–2020)

Final Draft

Version 3.0

March 30, 2020

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 childhood obesity data solutions. However, the CODI Implementation Guide does not represent official views or guidance of CDC.

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

© 2020, The MITRE Corporation. All Rights Reserved.

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.

Table of Contents

1. Introduction	1
1.1 Background	1
1.2 Purpose.....	2
1.3 Scope.....	2
1.4 Audience	3
1.5 Document Organization	3
2. CODI Background.....	4
2.1 CODI Data Models	4
2.2 CODI Roles.....	5
3. General Guidance	7
3.1 Data Cleaning Expectations	7
3.2 Missing Data	7
3.3 Reference Tables.....	7
3.4 Start Date	8
3.5 Community Health Partner Data.....	8
4. CDM Data Tables	9
4.1 DEMOGRAPHIC	9
4.2 ENCOUNTER	9
5. CODI Ancillary Tables	10
5.1 ALERT.....	10
5.2 ASSET_DELIVERY	10
5.3 FAMILY_HISTORY	11
5.4 IDENTIFIER.....	11
5.5 IDENTITY_HASH_BUNDLE.....	11
5.6 LINK.....	11
5.7 PROGRAM.....	12
5.8 REFERRAL	12
5.9 SESSION	13
5.10 SESSION_ALERT	14
6. Additional Resources.....	15
6.1 Requesting Changes.....	15
6.2 Questions.....	15
Appendix A. Additional Guidance for the Denver Pilot.....	16
Appendix B. RDM Documentation.....	17
B.1 Overview	17

B.2	Data Dictionary	21
B.2.1	ALERT	21
B.2.2	ASSET_DELIVERY	21
B.2.3	CENSUS_DEMOG	22
B.2.4	CENSUS_LOCATION	23
B.2.5	COST	24
B.2.6	CURRICULUM_COMPONENT	25
B.2.7	DEMOGRAPHIC	27
B.2.8	DIAGNOSIS	28
B.2.9	ENCOUNTER	29
B.2.10	FAMILY_HISTORY	30
B.2.11	LAB_RESULT_CM	31
B.2.12	LINK	32
B.2.13	PRESCRIBING	32
B.2.14	PROCEDURES	33
B.2.15	PROGRAM	34
B.2.16	PROVIDER	36
B.2.17	REFERRAL	37
B.2.18	SESSION	38
B.2.19	SESSION_ALERT	39
B.2.20	VITAL	39
B.3	CODI Research Data Model Codeset Details	40
B.3.1	ABN_TYPE	40
B.3.2	ADMITTING_SOURCE_TYPE	41
B.3.3	ASSET_TYPE	42
B.3.4	DIRECTION_TYPE	42
B.3.5	DISCHARGE_STATUS_TYPE	43
B.3.6	DISCHARGE_TYPE	43
B.3.7	DRG_TYPE	44
B.3.8	DX_SOURCE_TYPE	44
B.3.9	DX_TYPE	45
B.3.10	ENCOUNTER_TYPE	45
B.3.11	FACILITY_TYPE	46
B.3.12	FREQ_TYPE	47
B.3.13	GEOLEVEL_TYPE	47
B.3.14	LANGUAGE_TYPE	47
B.3.15	MODE_TYPE	48
B.3.16	MODIFIER_TYPE	48
B.3.17	ORGANIZATION_TYPE	49
B.3.18	ORIGIN_TYPE	49
B.3.19	PAYER_TYPE	49
B.3.20	PROCESS_PERFORMED_TYPE	49
B.3.21	PX_TYPE	50
B.3.22	QUAL_TYPE	50
B.3.23	RACE_TYPE	50
B.3.24	REFERRAL_STATUS_TYPE	51

B.3.25RELATIONSHIP_TYPE	51
B.3.26RESULT_LOC_TYPE	51
B.3.27RUCA_CODE_TYPE	52
B.3.28RX_ORIGIN_TYPE	53
B.3.29SETTING_TYPE	54
B.3.30SEX_TYPE	54
B.3.31SPECIALTY_TYPE	54
B.3.32SPECIMEN_SOURCE_TYPE	55
B.3.33UNIT_TYPE	55
B.3.34VITAL_SOURCE_TYPE	55
B.3.35YES//NO	55
Appendix C. RLDM Documentation.....	57
C.1 Overview	57
C.2 IDENTIFIER.....	58
C.3 IDENTITY_HASH_BUNDLE.....	59
Acronyms	60
List of References	Error! Bookmark not defined.
NOTICE	62

List of Figures

Figure 1: Overview of CODI Research and Record Linkage Data Models	2
Figure 2: Six Data Partners with a Common Data Coordinating Center	5
Figure 3: Example of Privacy Preserving Record Linkage Performed by a Linkage Agent.....	6
Figure 4: Sample Program and Session Data	13
Figure 5. RDM Tables from CDM, VDW, and OMOP	19
Figure 6. RDM Ancillary Tables.....	20
Figure 7. CODI Record Linkage Ancillary Tables	57

List of Tables

Table 1. Document Organization	3
Table 2: Conceptual Components of the RDM	4
Table 3. ALERT Details.....	21
Table 4. ASSET_DELIVERY Details	21
Table 5. CENSUS_DEMOG Details.....	22
Table 6. CENSUS_LOCATION Details.....	23
Table 7. COST Details	25
Table 8. CURRICULUM_COMPONENT Details	26
Table 9. DEMOGRAPHIC Details	27
Table 10. DIAGNOSIS Details	28
Table 11. ENCOUNTER Details	29
Table 12. FAMILY_HISTORY Details	30
Table 13. LAB_RESULT_CM Details	31
Table 14. LINK Details	32
Table 15. PRESCRIBING Details.....	33
Table 16. PROCEDURES Details.....	34
Table 17. PROGRAM Details	34
Table 18. PROVIDER Details.....	36
Table 19. REFERRAL Details	37

Table 20. SESSION Details	38
Table 21. SESSION_ALERT Details.....	39
Table 22. VITAL Details.....	40
Table 23. ABN_TYPE Details	40
Table 24. ADMITTING_SOURCE_TYPE Details	41
Table 25. ASSET_TYPE Details	42
Table 26. DIRECTION_TYPE Details	42
Table 27. DISCHARGE_STATUS_TYPE Details.....	43
Table 28. DISCHARGE_TYPE Details.....	44
Table 29. DRG_TYPE Details	44
Table 30. DX_SOURCE_TYPE Details	44
Table 31. DX_TYPE Details.....	45
Table 32. ENCOUNTER_TYPE Details	45
Table 33. FREQ_TYPE Details	47
Table 34. GEOLEVEL_TYPE Details.....	47
Table 35. LANGUAGE_TYPE Details	48
Table 36. MODE_TYPE Details.....	48
Table 37. MODIFIER_TYPE Details	48
Table 38. ORIGIN_TYPE Details.....	49
Table 39. PROCESS_PERFORMED_TYPE Details	49
Table 40. PX_TYPE Details.....	50
Table 41. RACE_TYPE Details	51
Table 42. REFERRAL_STATUS_TYPE Details	51
Table 43. RESULT_LOC_TYPE Details.....	52
Table 44. RUCA_CODE_TYPE Details.....	52
Table 45. RX_ORIGIN_TYPE Details	53
Table 46. SETTING_TYPE Details	54
Table 47. SEX_TYPE Details	54
Table 48. VITAL_SOURCE_TYPE Details	55
Table 49. YES//NO Details	55
Table 50. IDENTIFIER Details.....	58
Table 51. IDENTITY_HASH_BUNDLE Details	59

1. Introduction

As part of the Centers for Disease Control and Prevention's (CDC) efforts to promote health, prevent disease, injury, and disability, and prepare for emerging health threats, the Division of Nutrition, Physical Activity, and Obesity; and Center for Surveillance, Epidemiology, and Laboratory Services partnered with the CMS Alliance to Modernize Healthcare federally funded research and development center (Health FFRDC) for the Childhood Obesity Data Initiative (CODI). CODI will expand the ability to capture, standardize, integrate, and query existing patient-level electronic health record (EHR) and community data via a common data model. CODI augments an existing standard data model with information about weight-related interventions and social factors.

This document describes how CODI data partners—those organizations that participate in CODI by hosting data—should interpret the CODI data tables. It assumes that data partners are generally familiar with the Patient Centered Outcomes Research Network's (PCORnet) Common Data Model (CDM).¹ It also provides best practices for addressing situations in which a data partner may identify multiple ways to populate the CODI data models using local data.

1.1 Background

CODI distributes the work needed to answer researchers' queries using PCORnet infrastructure. CODI builds on the CDM by introducing ancillary tables, which are a way to introduce new functionality into a PCORnet or PCORnet-compatible clinical data network. CODI introduces ancillary tables drawn from the Colorado Health Observation Regional Data Service (CHORDS) virtual data warehouse (VDW) and the Observational Medical Outcomes Partnership (OMOP) CDM. CODI also introduces ancillary tables developed specifically for CODI.

The CODI Data Model comprises two distinct data models: 1) the CODI Research Data Model (RDM) introduces seven ancillary tables intended to supply data needed to answer researchers' queries; 2) the CODI Record Linkage Data Model (RLDM) introduces two ancillary tables used by CODI to match patients from different data partners. This linkage is used to construct a longitudinal patient record from the information supplied by the data partners. Figure 1 illustrates the major components of the CODI Data Model and the provenance for those components.

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.

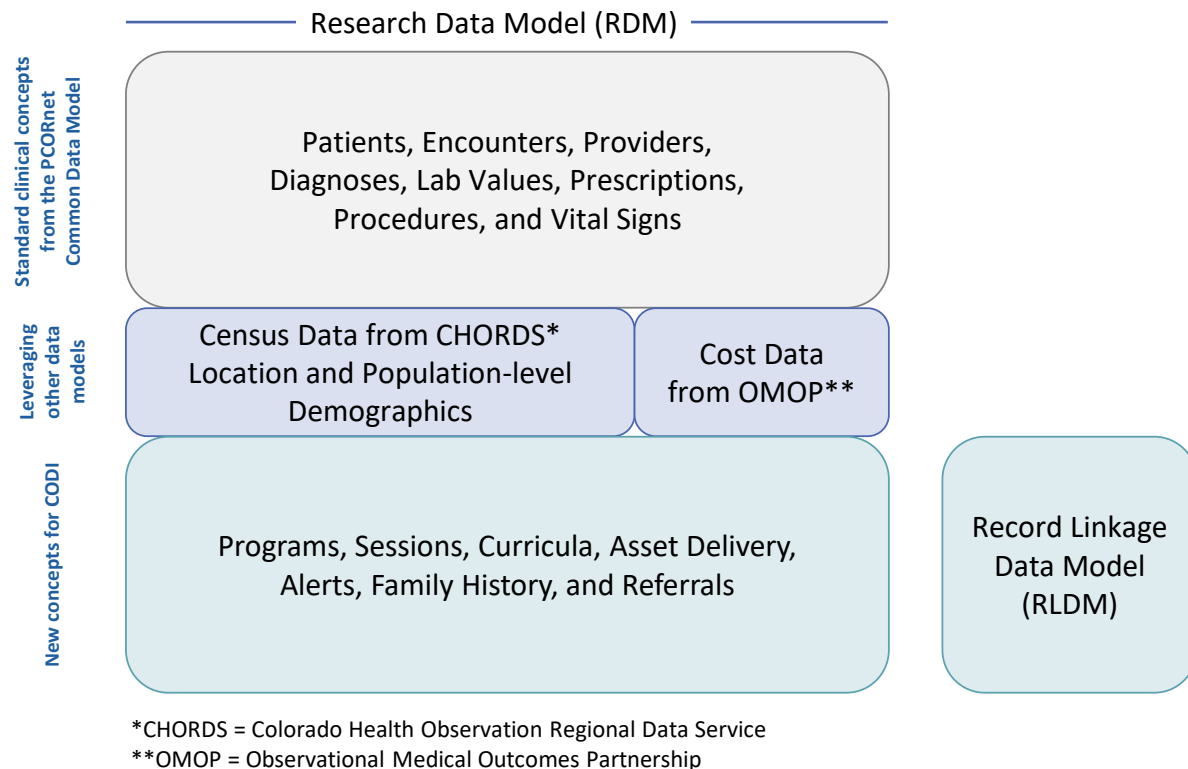


Figure 1: Overview of CODI Research and Record Linkage Data Models

This document is based upon the CODI Data Architecture Gaps and Recommendations report, which was informed by the research question formulation, and the decision by the CODI Collaborative Work Group (CCWG) during an in-person meeting in December 2019 to adopt privacy-preserving record linkage (PPRL) and a logical data warehouse query architecture. These decisions influence the guidance provided herein.

1.2 Purpose

The purpose of this document is to provide the implementation guidance necessary for data partners to implement the CODI RDM and CODI RLDM. Toward that end, this document provides:

- Descriptions of the RDM and RLDM
- General expectations regarding how to implement the tables within the CODI Data Models
- Specific guidance regarding individual data elements of the CODI Data Models

1.3 Scope

This guide provides implementation guidance for the ancillary tables unique to CODI and for the data elements introduced by CODI to tables incorporated from other sources (i.e., the lowest layer in Figure 1). This guide does not provide guidance for data tables documented elsewhere

(the top two layers in Figure 1); that implementation guidance can be found in the following documents:

- PCORnet Common Data Model v5.1 Specification²
- CHORDS VDW 3.1 Data Model Manual²
- [OMOP Common Data Model v6.0 Specifications](#)

The guidance provided in this document addresses all tables in the CODI Data Models. However, because implementation of the full RDM may be challenging initially, this guide indicates the priority for each data table in the RDM.

1.4 Audience

The primary audience for this document is the technical staff of the data partners' organizations—those individuals directly responsible for populating the CODI Data Models using their systems' data. The secondary audience includes project staff indirectly responsible for implementation and potential new data partners trying to assess the feasibility of implementing the CODI Data Models. Obesity researchers are likely most interested in the data model documentation appearing in the appendices or as a standalone data dictionary.

1.5 Document Organization

This document is organized as follows:

Table 1. Document Organization

Section	Purpose
Section 2: CODI Background	Provides an overview of the CODI data models and introduces common definitions
Section 3: General Guidance	Provides general guidance for data partners
Section 4: CDM Data Tables	Provides guidance on how to implement the changes made by CODI to existing CDM tables
Section 5: CODI Ancillary Tables	Provides guidance on how to implement the CODI ancillary tables
Appendix A: Additional Guidance for the Denver Pilot	Provides specific guidance for the Denver pilot
Appendix B: RDM Documentation	Presents the RDM conceptual data model
Appendix C: RLDM Documentation	Presents the detailed RLDM data dictionary
Acronym List	Defines the acronyms used in this document
List of References	Lists the sources used in preparing this document

² This document is available by sending a request to CODI@cdc.gov.

2. CODI Background

This section first summarizes the CODI Data Models. It then defines several roles within CODI relevant to implementing the CODI Data Models.

2.1 CODI Data Models

The RDM provides the data tables and data elements needed to answer selected critical childhood obesity research questions. Table 2 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.

Table 2: Conceptual Components of the RDM

Information about	Can be found in	Priority	Description
The child	DEMOGRAPHIC	High	Includes demographic information about the child
Risk factors	FAMILY HISTORY	Medium	Includes details about any family members' health conditions
Clinical care	ENCOUNTER DIAGNOSIS LAB VALUE PRESCRIPTION PROCEDURE PROVIDER VITAL	High High High High High High High	Includes information about the child's interactions with the healthcare delivery system
Weight-related interventions	PROGRAM CURRICULUM COMPONENT SESSION ASSET DELIVERY	High Medium High Medium	Includes details about intervention aims and settings (PROGRAM); how the interventions are structured (CURRICULUM COMPONENT); who is administering the intervention and how (SESSION); if an asset (e.g., food, money) was provided (ASSET DELIVERY)
Referrals	REFERRAL	Medium	Includes incoming and outgoing referrals within and across organizations
Clinical decision support	ALERT SESSION ALERT	Medium Medium	Includes details about the types of clinical alerts (ALERT) and when they triggered (SESSION ALERT)
Cost of care	COST	Low	Includes information about the amounts charged
Child's location	CENSUS LOCATION CENSUS DEMOG	High High	Includes details about where the child has resided (CENSUS LOCATION) and population-level demographic data about that location from the census (CENSUS DEMOG)

The RLDM provides the data tables and data elements needed to perform the record linkage process. It includes two tables: IDENTIFIER contains unencrypted personally identifiable information (PII); IDENTITY_HASH_BUNDLE contains corresponding encrypted personally identifiable information. The manner in which the IDENTITY_HASH_BUNDLE table is populated using PII will be described in the forthcoming CODI Privacy Preserving Record Linkage Implementation Guide.

2.2 CODI Roles

A *data partner* is an organization that participates in CODI by hosting data. Many data partners will host their organization's own data. In other cases (e.g., community health partners), the organization that contributes data will rely on an intermediary data partner to host their data.

A *data coordinating center* is an organization with several responsibilities:

- It distributes research queries to data partners.
- It assembles the results from data partners into longitudinal records.
- It distributes references tables to data partners.

Figure 2 shows how researchers interact with the data coordinating center, which distributes their research queries to data partners. The data coordinating center assembles the results into longitudinal records, which are sent to the researchers.

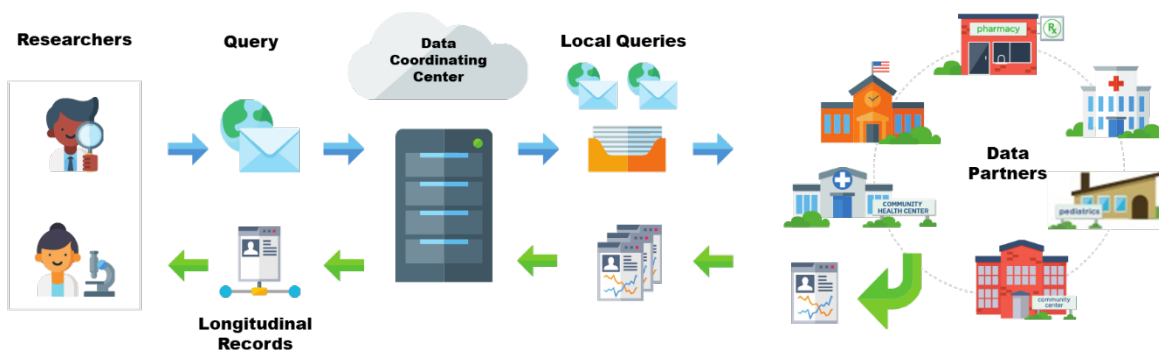


Figure 2: Six Data Partners with a Common Data Coordinating Center

A *linkage agent* is an organization that performs linkage on behalf of data partners. The linkage agent receives encrypted PII and produces globally unique LINKIDs used to construct longitudinal records. Figure 3 illustrates a linkage agent receiving hashed data from data partners, which the linkage agent uses to link children across those data partners.

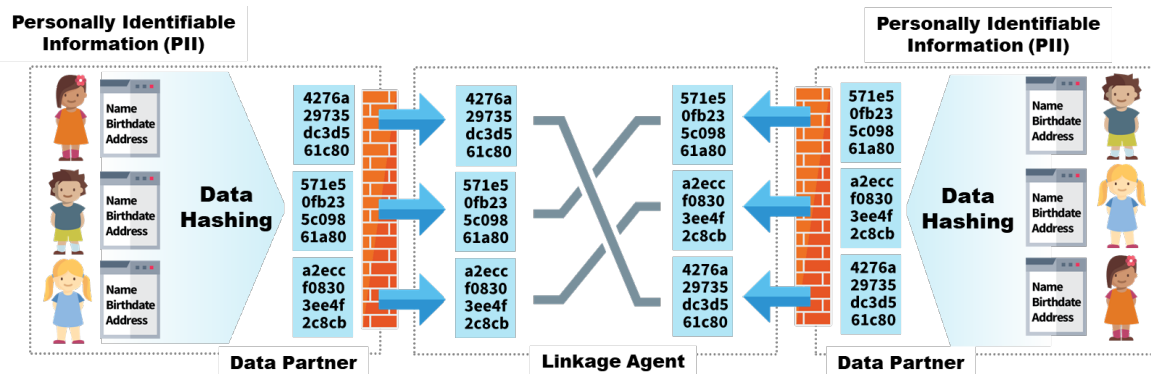


Figure 3: Example of Privacy Preserving Record Linkage Performed by a Linkage Agent

An *implementing network* is a collection of data partners that share a common linkage agent and data coordinating center. For example, the initial pilot demonstration is a single implementing network based in the Denver, Colorado area.

3. General Guidance

This section provides general guidance for data partners. 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, data partners 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. The intent is that any data cleaning or inferences will be performed post-hoc by researchers based on their research 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 partner can determine which process steps necessarily follow from observations recorded in their systems.

Mapping from a local codeset to a CODI codeset is anticipated and does not constitute data cleaning. I.e., data partners *should* map their codes to CODI.

3.2 Missing Data

CDM uses the [HL7 conventions](#) 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.
- If the source value is null or blank, CDM uses NI (no information).
- If the source value is an explicit unknown value, CDM uses UN (unknown).
- When the source value cannot be mapped to CDM, CDM uses OT (other).

3.3 Reference Tables

With the RDM, there are three tables likely to not be populated from an EHR or another IT system. These include ALERT, CENSUS_DEMOG, and PROGRAM. The CENSUS_DEMOG table contains information from the Census Bureau. The data coordinating center (DCC) will share this file with data partners using PopMedNet or a similar mechanism agreed upon with the data partners.

The remaining two tables must be populated manually. Data partners are encouraged to populate these tables with data provided manually as part of the extract–transform–load (ETL) process that populates the remainder of the RDM. For example, these tables might be populated using a series of SQL INSERT statements using explicit VALUES instead of a SELECT statement. Data partners 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 SESSION and SESSION_ALERT).

3.4 Start Date

Each implementing network should establish a *start date* for that network. The start date represents the earliest possible date for which data partners can reliably populate the CODI Data Models. The start date is used in Sections 4 and 5 to describe how to implement tables in the CODI Data Models.

3.5 Community Health Partner Data

Within the RDM, only the tables listed below are relevant³ to community health partners. Community data partners should create the remaining RDM tables in their data warehouses but leave those tables empty so queries that reference those tables do not fail. *Italics* indicate lower priority tables. (All tables in the RLDM are relevant to community health partners.)

- *ASSET_DELIVERY*
- CENSUS_DEMOG
- CENSUS_LOCATION
- *COST*
- *CURRICULUM_COMPONENT*
- DEMOGRAPHIC
- PROGRAM
- PROVIDER
- *REFERRAL*
- SESSION

³ Additional data tables may become relevant to community health partners after the initial pilot demonstration. For example, some community health partners might capture VITAL signs, such as height and weight.

4. CDM Data Tables

This section provides implementation guidance for the CDM tables to which CODI adds additional data elements. Additional implementation guidance for these tables is available in the [PCORnet Common Data Model v5.1 Specification](#), and an enumeration of the attributes of these tables appears in Appendix B. The remainder of this section provides guidance specific to the concepts added to the CDM tables for CODI. Section 5 provides implementation guidance for each of the CODI-specific ancillary tables. Within both sections, data tables and attributes appear in ALL CAPS; see Appendix B and Appendix C for definitions of these attributes.

4.1 DEMOGRAPHIC

The DEMOGRAPHIC table contains a single record for each patient. This table should include information for children (and only children) between the ages of 2 and 19 (inclusive) with at least one clinical visit or program participation since the implementing network's start date. Implementers should not include children without other records in the RDM. For example, a child should be included in the DEMOGRAPHIC table if they have ENCOUNTER data, but not VITAL data.

To preserve referential integrity, there must be a DEMOGRAPHIC record for any child 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 children without other information in the RDM because populating the DEMOGRAPHIC table with all children introduces the possibility that their PII is shared with the DCC even though insufficient information exists about those children to answer possible research questions. For example, a child 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 child.

4.2 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, LAB_RESULT_CM, PRESCRIBING, PROCEDURES, and VITAL. These references are optional so that diagnoses, 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 not known. In some cases, there may not even be a corresponding encounter, such as when vital signs are measured outside a clinical setting.

5. CODI Ancillary Tables

5.1 ALERT

The ALERT table contains one record for each distinct *kind* of alert. This table should only be populated with alerts directly related to obesity, which each data partner will need to determine on behalf of their organization. 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](#) (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.

5.2 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. 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 of time during which these benefits were received, with a DELIVERY_FREQ_UNIT of Monthly and a DELIVERY_FREQ of 2.

Data partners 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.

5.3 CURRICULUM_COMPONENT

The CURRICULUM_COMPONENT table enumerates the standard elements of a program. It supports 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.

5.4 FAMILY_HISTORY

The FAMILY_HISTORY table stores information regarding a child's family health. Each entry records a single condition reported for a family member. Thus, if a child's parents both have a history of obesity, two records would be present in this table. The intention is that data partners only retrieve family history information present in the child's 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 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.

5.5 IDENTIFIER

The IDENTIFIER table contains one row for each unique combination of identifying information. For example, if a data partner captures the names of two parents for each child, then the IDENTIFIER table would contain two rows for each child. These rows would have the same data for the child's information and differing data for the parents' information. If a data partner has a current address and two former addresses for a child, that would result in three rows in the IDENTIFIER table. If both scenarios apply, then the data partner would generate six rows in the IDENTIFIER table (two parents times three addresses). More generally, the IDENTIFIER table contains the cross-product of all possible identifying characteristics.

Every child appearing in the DEMOGRAPHIC table should appear at least once in the IDENTIFIER table. Similarly, every child in the IDENTIFIER table should also appear (once) in the DEMOGRAPHIC table.

5.6 IDENTITY_HASH_BUNDLE

The IDENTITY_HASH_BUNDLE table contains one record for each record in the IDENTIFIER table. The latter contains unencrypted PII. The former contains encrypted versions of the PII generated using a [cryptographic hash function](#). The CODI Privacy Preserving Record Linkage Implementation Guide⁴ will describe the specific procedure used to generate hash bundles from PII, i.e., the transformation that data partners will need to apply to records in the IDENTIFIER table to generate records in the IDENTITY_HASH_BUNDLE table. The deidentified data in this table are segregated from the PII in the IDENTIFIER table to simplify describing which data are shared with a linkage agent as part of the record linkage process. The PII in the IDENTIFIER table are not shared. The deidentified data in this table are shared.

5.7 LINK

The LINK table contains one record for each person in the demographics table for each iteration of record linkage. When the RDM 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 each data partner with a partner- and iteration- specific mapping from PATID to LINKID. Upon receipt of this mapping, data partners should populate the LINK table for that iteration.

⁴ This guide is still under development.

5.8 PROGRAM

The PROGRAM table contains one record for each distinct weight-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 weight-related program 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 ten 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. 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 child'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 parent program with which they are affiliated.

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). Data partners should provide all three variants for each program, if possible.

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 mission session information.

There are no dates of enrollment or completion associated with a program for two reasons. First, the program table describes how the program is administered irrespective of any child's participation in the program. Enrollment and completion dates would need to be stored in a separate program participation table. Such a table does not exist because the Technical Environmental Scan determined that enrollment is often hard to distinguish from attendance (i.e., the first session attended indicates enrollment). Completion date was almost never available. A researcher interested in program completion might compare the cumulative dose received with the prescribed total dose for that program.

5.9 REFERRAL

The REFERRAL table contains one record for each outgoing or incoming referral. The DIRECTION attribute indicates if the record represents a data partner 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 a child with a weight-related program. Implementers will need to map organizations to CMS Certification Numbers; see ORGANIZATION_TYPE in Appendix B for more information about coding organizations.

5.10 SESSION

The SESSION table contains one record for each interaction between a child 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, exercise and nutrition screening may transpire. When an ENCOUNTER involves multiple providers interacting with a child, multiple SESSION records should be created. For example, a single encounter sometimes includes a child 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 4. 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 that child 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.

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

Figure 4: Sample Program and Session Data

At a minimum, data partners should populate the SESSION table for well-child visits and for encounters that are part of a weight-related program (e.g., 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 obesity or obesity-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, data partners would populate the SESSION table for every encounter.

If a data partner partially implements the SESSION table, no record should be created for encounter types that have not been mapped to the SESSION table. For example, if a data partner implements the SESSION table for only well-child visits and weight-related program encounters, then no record in the SESSION table should be created for emergency department encounters.

For weight-related programs based in a community setting, each SESSION record corresponds to a child's participation in the program. For example, a child that completes a program that meets weekly ten times should have ten distinct SESSION records.

The DOSE attribute indicates the amount of time spent interacting with the child (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.

5.11 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 the provider responsible for a SESSION was made aware of a given ALERT.

6. Additional Resources

6.1 Requesting Changes

Data partners and researcher are likely to find limitations with the Implementation Guide or CODI Data Models. The following process will be followed to process those change requests.

1. The data partner or researcher documents the change request and sends that request to CODI@cdc.gov.
2. MITRE determines which documents or data models, if any, might need to be changed.
3. MITRE presents the potential changes to the Research Question subgroup (within two weeks of the initial request) for feedback.
4. CDC decides how to handle the request based on the subgroup feedback.
5. MITRE implements any necessary changes and uploads the new documents to Basecamp.
6. Basecamp notifies subscribers of the availability of the updated documents.

Changes will generally increment the minor version number of the Implementation Guide (e.g., from 1.4 to 1.5). Every six months, CDC and MITRE will increment the major version number after completing the necessary clearance processes.

6.2 Questions

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

Appendix A. Additional Guidance for the Denver Pilot

- The DCC for the Denver pilot will be the University of Colorado.
- The start date in Section 3.4 for the Denver pilot is May 1, 2016. The RDM should be populated with information subsequent to that 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 is necessarily incomplete.
- The Denver pilot anticipates completing the record linkage process yearly.

Appendix B. RDM Documentation

This appendix comprises three sections. The first section provides an overview of the RDM in its entirety. The second section provides details for each data table within the RDM. The third section provides details for each codeset within the RDM.

B.1 Overview

The RDM is based primarily on CDM. It also incorporates additional census and location tables from VDW and an additional cost table from OMOP (Figure 5). Finally, the RDM introduces seven new ancillary tables (

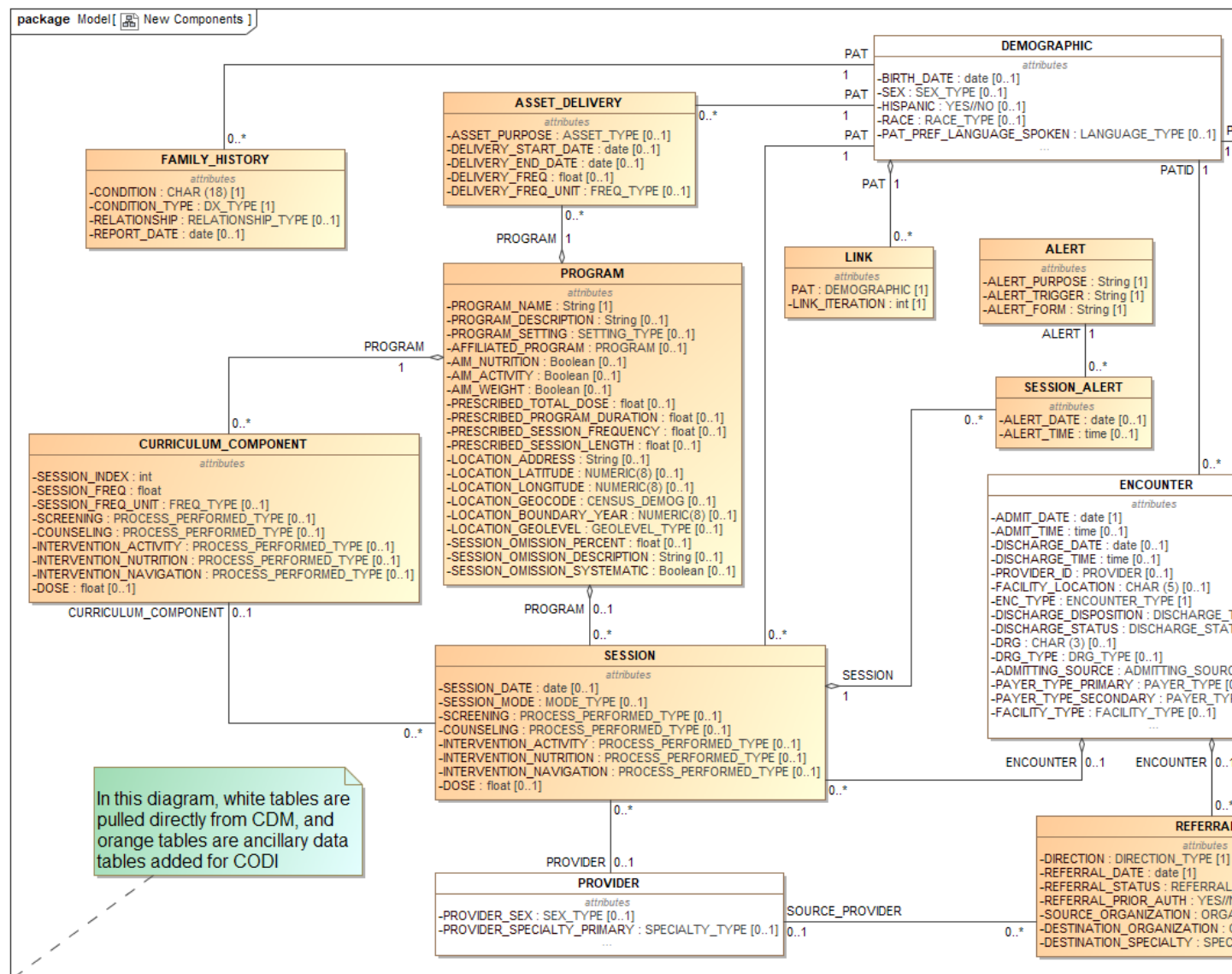


Figure 6).

Within these figures:

- Color indicates the source of the table.
 - White tables are from CDM.
 - Yellow tables are from VDW or OMOP.
 - Orange tables are new ancillary tables specific to CODI.
- Lines indicate primary key / foreign key relationships among the tables.
 - Labels in ALL CAPS are the column names of the foreign key (e.g., PATID is a column in the ENCOUNTER table that is a foreign key that references the DEMOGRAPHIC table).
 - 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 and can be interpreted the same as those in ALL CAPS.
- Numbers 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).

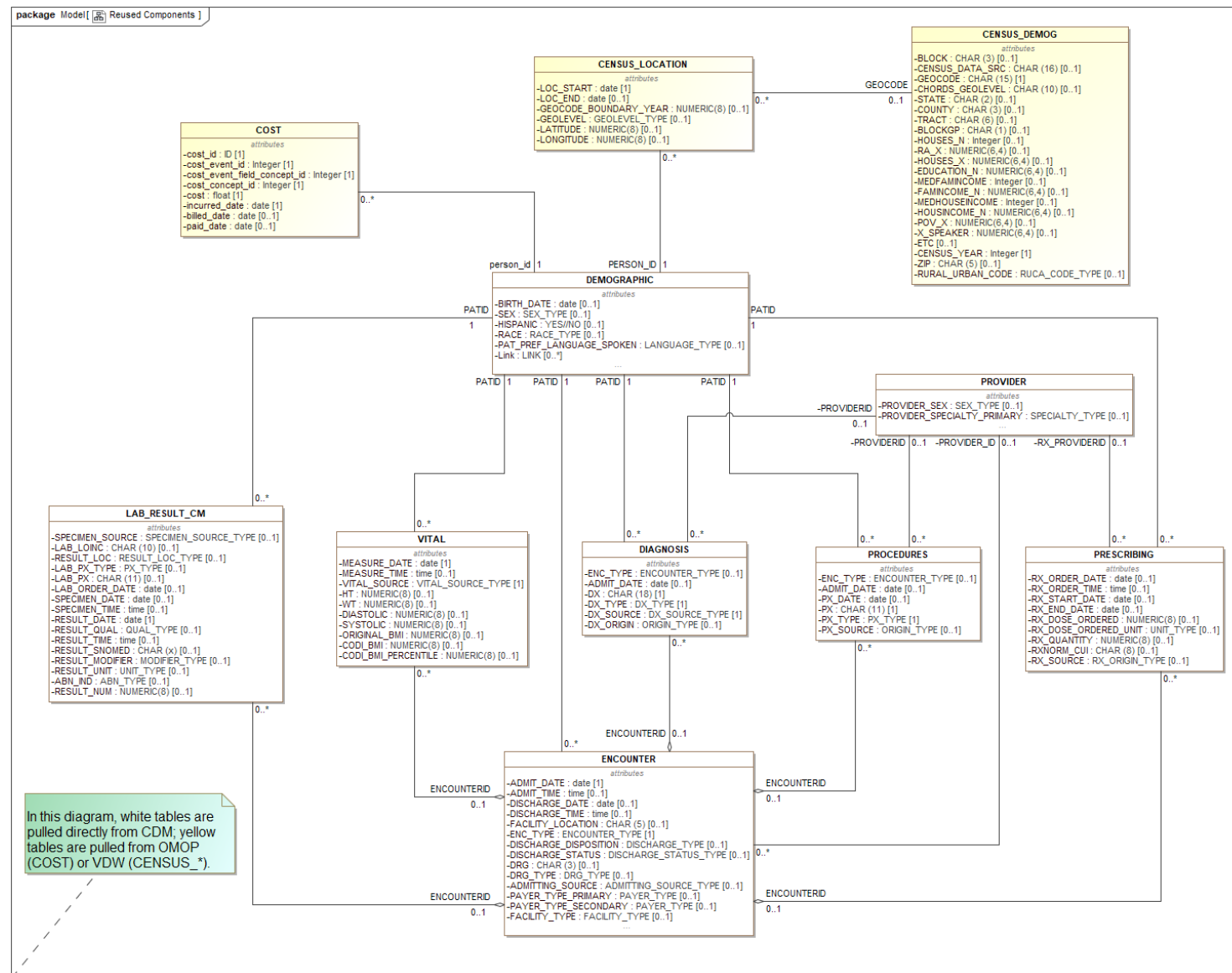


Figure 5. RDM Tables from CDM, VDW, and OMOP

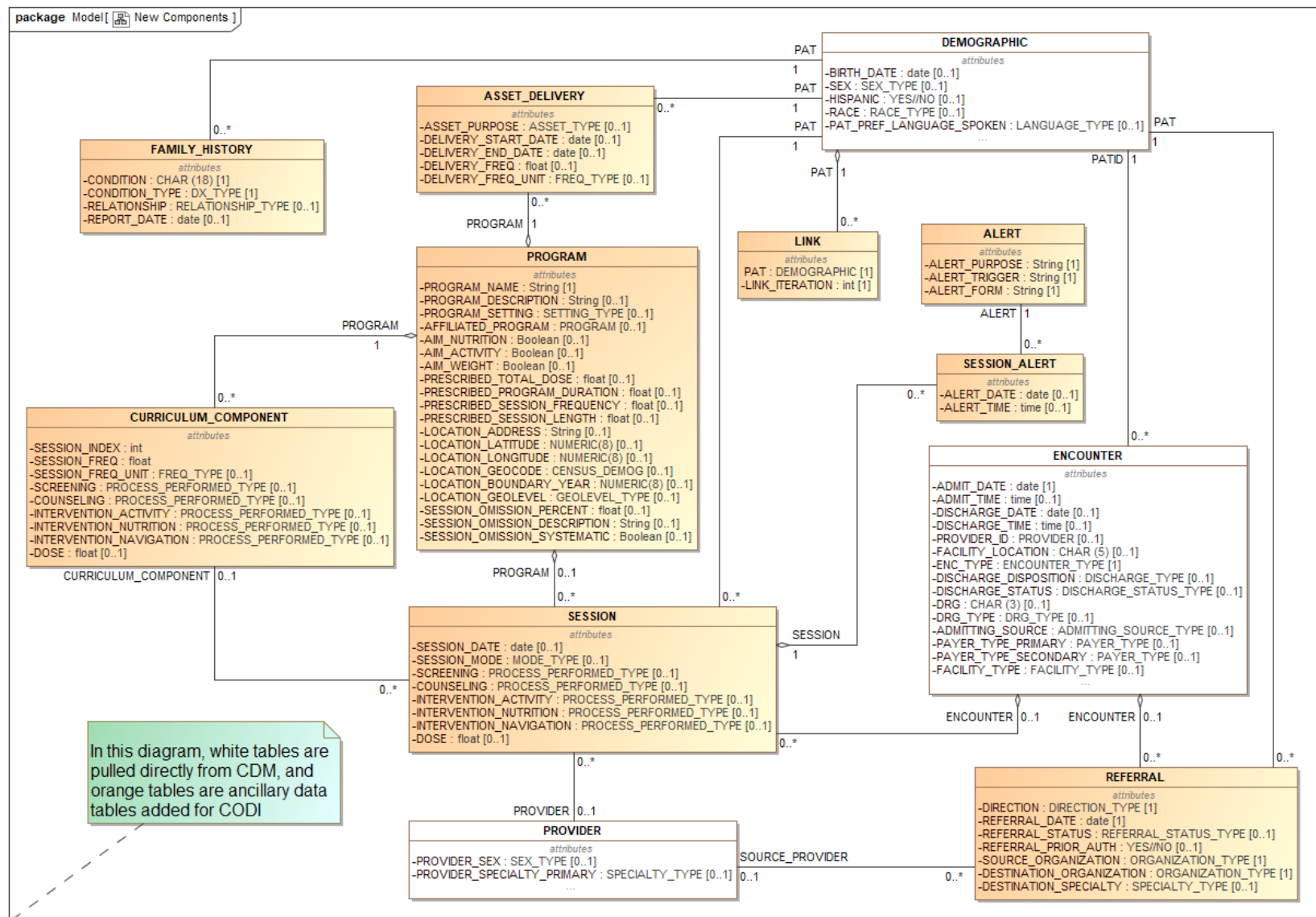


Figure 6. RDM Ancillary Tables

B.2 Data Dictionary

B.2.1 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 3. ALERT Details

Attribute	Cardinality	Type	Definition
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.

B.2.2 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 4. ASSET_DELIVERY Details

Attribute	Cardinality	Type	Definition
ASSET_DELIVERY_ID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographic table.
PROGRAMID	1	FK::PROGRAM	A link back to the program table.

Attribute	Cardinality	Type	Definition
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.

B.2.3 CENSUS_DEMOG

The [VDW] CENSUS_DEMOG table is a static reference table that will be provided to each data partner through the PopMedNet data sharing client. This data is used to provide community level attributes for each census tract or county of a patient's residence.

For complete documentation, see the CHORDS VDW Data Model Manual.

Table 5. CENSUS_DEMOG Details

Attribute	Cardinality	Type	Definition
BLOCK	0..1	CHAR (3)	A three-character numeric code assigned to census blocks.
CENSUS_DATA_SRC	0..1	CHAR (16)	Source of the data in the record
GEOCODE	1	CHAR (15)	A numeric code signifying the geographic location of the address where the person was found.
CHORDS_GEOLEVEL	0..1	CHAR (10)	[Indicates the specificity of the GEOCODE. It is the same variable as CENSUS_LOCATION::GEOLEVEL using a different value set: Block, BlkGrp, or Tract.]
STATE	0..1	CHAR (2)	A two-character numeric code assigned to US states, districts, territories and protectorates.
COUNTY	0..1	CHAR (3)	A three-character numeric code assigned to census counties.
TRACT	0..1	CHAR (6)	A six-character numeric code assigned to census tracts.
BLOCKGP	0..1	CHAR (1)	A one-character numeric code assigned to census block groups.
HOUSES_N	0..1	Integer	Number of housing units in geography

Attribute	Cardinality	Type	Definition
RA_X	0..1	NUMERIC(6,4)	Percent [Race X]
HOUSES_X	0..1	NUMERIC(6,4)	Proportion of occupied housing units that are X.
EDUCATION_N	0..1	NUMERIC(6,4)	[Education level]
MEDFAMINCOME	0..1	Integer	Median family income
FAMINCOME_N	0..1	NUMERIC(6,4)	[Income level]
MEDHOUSEINCOME	0..1	Integer	Median household income
HOUSINCOME_N	0..1	NUMERIC(6,4)	[Income level]
POV_X	0..1	NUMERIC(6,4)	[Poverty range]
X_SPEAKER	0..1	NUMERIC(6,4)	Proportion of people over age 5 that speak only X or speak X "very well"
ETC	0..1		[For complete documentation, see the CHORDS VDW Data Model Manual.]
CENSUS_YEAR	1	Integer	Year the census data was collected
ZIP	0..1	CHAR (5)	A five-character numeric code assigned by the US Postal Service to various regions where it delivers mail.
RURAL_URBAN_CODE	0..1	RUCA_CODE_T YPE	A classification of the population density, urbanization, and commuting patterns of the census region.
CensusLocation	0..*	FK::CENSUS_LO CATION	A set of patient locations that correspond to this census area.

B.2.4 CENSUS_LOCATION

The VDW CENSUS_LOCATION table holds patient geographic location information collected at healthcare encounters. Patient addresses should be geocoded, and FIPS codes down to the census tract level should be populated in the CENSUS_LOCATION table.

Table 6. CENSUS_LOCATION Details

Attribute	Cardinality	Type	Definition
PERSON_ID	1	FK::DEMOGRAP HIC	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.

Attribute	Cardinality	Type	Definition
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.
GEOCODE	0..1	FK::CENSUS_DEMOG	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)	The latitude of the location.
LONGITUDE	0..1	NUMERIC(8)	The longitude of the location.

B.2.5 COST

The OMOP 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 7. COST Details

Attribute	Cardinality	Type	Definition
cost_id	1	ID	A unique identifier for each COST record.
person_id	1	FK::DEMOGRAP HIC	A unique identifier for each PERSON.
cost_event_id	1	Integer	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	Integer	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	Integer	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

B.2.6 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. Or, a program can comprise a recurring curriculum with no endpoint and a set of standard sessions that recur with some frequency.

Table 8. CURRICULUM_COMPONENT Details

Attribute	Cardinality	Type	Definition
CURRICULUM_COMPONENT_ID	1	ID	
PROGRAMID	1	FK::PROGRAM	A link back to the program this component of a curriculum belongs to.
SESSION_INDEX		int	An ordinal used to establish a total ordering on the sessions within a fixed curriculum.
SESSION_FREQ		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.
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.

B.2.7 DEMOGRAPHIC

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

Table 9. DEMOGRAPHIC Details

Attribute	Cardinality	Type	Definition
PATID	1	ID	
BIRTH_DATE	0..1	date	Date of birth.
SEX	0..1	SEX_TYPE	Sex assigned at birth.
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.
ClearTextIdentifier	1..*	FK::IDENTIFIER	A set of clear-text identifiers associated with this person.
HashedIdentifier	1..*	FK::IDENTITY_HASH_BUNDLE	A set of hashed identifiers associated with this person.
AssetDelivery	0..*	FK::ASSET_DELIVERY	A set of asset deliveries associated with this person.
CensusLocation	0..*	FK::CENSUS_LOCATION	A set of census locations associated with this person.
Cost	0..*	FK::COST	A set of charges associated with this person.
Diagnosis	0..*	FK::DIAGNOSIS	A set of diagnoses associated with this person.
Encounter	0..*	FK::ENCOUNTER	A set of encounters associated with this person.
FamilyHistory	0..*	FK::FAMILY_HISTORY	A set of family histories associated with this person.
LabResult	0..*	FK::LAB_RESULT_CM	A set of lab results associated with this person.

Attribute	Cardinality	Type	Definition
Prescribing	0..*	FK::PRESCRIBING	A set of prescriptions associated with this person.
Procedure	0..*	FK::PROCEDURES	A set of procedures 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.
Link	0..*	FK::LINK	A set of record linkage identifiers associated with this person.

B.2.8 DIAGNOSIS

The CDM 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 10. DIAGNOSIS Details

Attribute	Cardinality	Type	Definition
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	CHAR (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.

Attribute	Cardinality	Type	Definition
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.

B.2.9 ENCOUNTER

The CDM 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 11. ENCOUNTER Details

Attribute	Cardinality	Type	Definition
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.
PROVIDER_ID	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).

Attribute	Cardinality	Type	Definition
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.
Diagnosis	0..*	FK::DIAGNOSIS	A set of diagnoses associated with this encounter.
LabResult	0..*	FK::LAB_RESULT_CM	A set of lab results 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.

B.2.10 FAMILY_HISTORY

The FAMILY_HISTORY table stores information regarding a child'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 12. FAMILY_HISTORY Details

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

Attribute	Cardinality	Type	Definition
CONDITION	1	CHAR (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.

B.2.11 LAB_RESULT_CM

The CDM 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 13. LAB_RESULT_CM Details

Attribute	Cardinality	Type	Definition
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	CHAR (10)	Use this field to store the LOINC code of the laboratory result.
RESULT_LOC	0..1	RESULT_LOCATION_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	CHAR (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.

Attribute	Cardinality	Type	Definition
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	CHAR (x)	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(8)	Standardized/converted result for quantitative results.

B.2.12 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 14. LINK Details

Attribute	Cardinality	Type	Definition
LINKID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographics table.
LINK_ITERATION	1	int	An iteration of the record linkage process.

B.2.13 PRESCRIBING

The CDM 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 15. PRESCRIBING Details

Attribute	Cardinality	Type	Definition
PRESCRIBINGID	1	ID	
PATID	1	FK::DEMOGRAP HIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTE R	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(8)	Dose of a given medication, as ordered by the provider
RX_DOSE_ORDERED_U NIT	0..1	UNIT_TYPE	Units of measure associated with the dose of the medication as ordered by the provider
RX_QUANTITY	0..1	NUMERIC(8)	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_TY PE	Source of the prescribing information.

B.2.14 PROCEDURES

The CDM 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 16. PROCEDURES Details

Attribute	Cardinality	Type	Definition
PROCEDURESID	1	ID	
PATID	1	FK::DEMOGRAP HIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTE R	Arbitrary encounter-level identifier. Used to link across tables.
ENC_TYPE	0..1	ENCOUNTER_T YPE	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.

B.2.15 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 17. PROGRAM Details

Attribute	Cardinality	Type	Definition
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.

Attribute	Cardinality	Type	Definition
PROGRAM_SETTING	0..1	SETTING_TYPE	A setting in which the program is offered (clinical or community).
AFFILIATED_PROGRAM_ID	0..1	FK::PROGRAM	A parent program of which this program is a component.
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) a child 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.
LOCATION_LATITUDE	0..1	NUMERIC(8)	A latitude of the corresponding address location.
LOCATION_LONGITUDE	0..1	NUMERIC(8)	A latitude of the corresponding address location.
LOCATION_GEOCODE	0..1	FK::CENSUS_DEMOG	A primary location at which this program's sessions are administered, expressed as a geocode.
LOCATION_BOUNDARY_YEAR	0..1	NUMERIC(8)	A census year for which the corresponding geocode location applies.
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.
Session	0..*	FK::SESSION	A set of sessions associated with this program.

B.2.16 PROVIDER

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

Table 18. PROVIDER Details

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

B.2.17 REFERRAL

The REFERRAL table contains one record for each outgoing or incoming referral.

Table 19. REFERRAL Details

Attribute	Cardinality	Type	Definition
REFERRALID	1	ID	
PATID	1	FK::DEMOGRAP HIC	A link back to the demographic table.
ENCOUNTERID	0..1	FK::ENCOUNTE R	A link back to the encounter table, if the referral can be unambiguously associated with an encounter.
DIRECTION	1	DIRECTION_TYP E	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_STA TUS_TYPE	A final disposition of the referral.
REFERRAL_PRIOR_AUT H	0..1	YES//NO	An indication of whether prior authorization was required for the referral.
SOURCE_PROVIDERID	0..1	FK::PROVIDER	A provider responsible for initiating this referral.
SOURCE_ORGANIZATIO N	1	ORGANIZATION _TYPE	An organization that initiated the referral.
DESTINATION_ORGANI ZATION	1	ORGANIZATION _TYPE	An organization to which the referral was sent.
DESTINATION_SPECIAL TY	0..1	SPECIALTY_TYP E	A clinical specialty for which the patient is being referred.

B.2.18 SESSION

The SESSION table contains one record for each session. A session is a specific point in time where a child/family is involved in programming that focuses on obesity, obesity prevention, healthy eating, or active living.

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 20. SESSION Details

Attribute	Cardinality	Type	Definition
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.
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.

Attribute	Cardinality	Type	Definition
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 a child 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.

B.2.19 SESSION_ALERT

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

Table 21. SESSION_ALERT Details

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

B.2.20 VITAL

The CDM 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 22. VITAL Details

Attribute	Cardinality	Type	Definition
VITALID	1	ID	
PATID	1	FK::DEMOGRAP HIC	Arbitrary person-level identifier. Used to link across tables.
ENCOUNTERID	0..1	FK::ENCOUNTE R	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(8)	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(8)	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(8)	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(8)	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(8)	BMI if calculated in the source system. Decimal precision is permissible.

B.3 CODI Research Data Model Codeset Details

B.3.1 ABN_TYPE

[From CDM]

Table 23. ABN_TYPE Details

Code	Definition
AB	Abnormal

Code	Definition
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

B.3.2 ADMITTING_SOURCE_TYPE

[From CDM]

Table 24. ADMITTING_SOURCE_TYPE Details

Code	Definition
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

Code	Definition
IH	Intra-hospital
NI	No information
UN	Unknown
OT	Other

B.3.3 ASSET_TYPE

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

Table 25. ASSET_TYPE Details

Code	Definition
CC	Child care
FO	Food
HI	Health insurance
TR	Transportation
NI	No information
UN	Unknown
OT	Other

B.3.4 DIRECTION_TYPE

This codeset enumerates the direction of a referral.

Table 26. DIRECTION_TYPE Details

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

B.3.5 DISCHARGE_STATUS_TYPE

[From CDM]

Table 27. DISCHARGE_STATUS_TYPE Details

Code	Definition
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

B.3.6 DISCHARGE_TYPE

[From CDM]

Table 28. DISCHARGE_TYPE Details

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

B.3.7 DRG_TYPE

[From CDM]

Table 29. DRG_TYPE Details

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

B.3.8 DX_SOURCE_TYPE

[From CDM]

Table 30. DX_SOURCE_TYPE Details

Code	Definition
AD	Admitting
DI	Discharge
FI	Final
IN	Interim

Code	Definition
NI	No information
UN	Unknown
OT	Other

B.3.9 DX_TYPE

[From CDM]

Table 31. DX_TYPE Details

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

B.3.10 ENCOUNTER_TYPE

[From CDM]

Table 32. ENCOUNTER_TYPE Details

Code	Definition
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

Code	Definition
	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

B.3.11 FACILITY_TYPE

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

B.3.12 FREQ_TYPE

This codeset enumerates different frequencies that something happens.

Table 33. FREQ_TYPE Details

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

B.3.13 GEOLEVEL_TYPE

[From VDW]

Table 34. GEOLEVEL_TYPE Details

Code	Definition
B	Block
G	Block Group
T	Census Tract
C	County
Z	Zip Code
P	Post Office
U	Unknown, unable to append

B.3.14 LANGUAGE_TYPE

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

Table 35. LANGUAGE_TYPE Details

Code	Definition
------	------------

B.3.15 MODE_TYPE

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

Table 36. MODE_TYPE Details

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

B.3.16 MODIFIER_TYPE

[From CDM]

Table 37. MODIFIER_TYPE Details

Code	Definition
EQ	Equal
GE	Greater than or equal to
GT	Greater than
LE	Less than or equal to
LT	Less than
TX	Text
NI	No information
UN	Unknown
OT	Other

B.3.17 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.

B.3.18 ORIGIN_TYPE

[From CDM]

Table 38. ORIGIN_TYPE Details

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

B.3.19 PAYER_TYPE

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

B.3.20 PROCESS_PERFORMED_TYPE

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

Table 39. PROCESS_PERFORMED_TYPE Details

Code	Definition
Y	Yes, the process step was conducted.

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

B.3.21 PX_TYPE

[From CDM]

Table 40. PX_TYPE Details

Code	Definition
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

B.3.22 QUAL_TYPE

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

B.3.23 RACE_TYPE

[From CDM]

Table 41. RACE_TYPE Details

Code	Definition
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

B.3.24 REFERRAL_STATUS_TYPE

[From CDM]

Table 42. REFERRAL_STATUS_TYPE Details

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

B.3.25 RELATIONSHIP_TYPESee <https://www.hl7.org/fhir/valueset-relatedperson-relationshiptype.html>**B.3.26 RESULT_LOC_TYPE**

[From CDM]

Table 43. RESULT_LOC_TYPE Details

Code	Definition
L	Lab
P	Point of Care
NI	No information
UN	Unknown
OT	Other

B.3.27 RUCA_CODE_TYPE

The rural-urban commuting area (RUCA) codes classify U.S. census tracts using measures of population density, urbanization, and daily commuting. The most recent RUCA codes are based on data from the 2010 decennial census and the 2006-10 American Community Survey. (See <https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx> for more information.)

Table 44. RUCA_CODE_TYPE Details

Code	Definition
1.0	Metropolitan area core: primary flow within an urbanized area (UA)—No additional code
1.1	Metropolitan area core: primary flow within an urbanized area (UA)—Secondary flow 30% to 50% to a larger UA
2.0	Metropolitan area high commuting: primary flow 30% or more to a UA—No additional code
2.1	Metropolitan area high commuting: primary flow 30% or more to a UA—Secondary flow 30% to 50% to a larger UA
3.0	Metropolitan area low commuting: primary flow 10% to 30% to a UA—No additional code
4.0	Micropolitan area core: primary flow within an urban cluster of 10,000 to 49,999 (large UC)—No additional code
4.1	Micropolitan area core: primary flow within an urban cluster of 10,000 to 49,999 (large UC)—Secondary flow 30% to 50% to a UA
5.0	Micropolitan high commuting: primary flow 30% or more to a large UC—No additional code
5.1	Micropolitan high commuting: primary flow 30% or more to a large UC—Secondary flow 30% to 50% to a UA
6.0	Micropolitan low commuting: primary flow 10% to 30% to a large UC—No additional code

Code	Definition
7.0	Small town core: primary flow within an urban cluster of 2,500 to 9,999 (small UC)—No additional code
7.1	Small town core: primary flow within an urban cluster of 2,500 to 9,999 (small UC)—Secondary flow 30% to 50% to a UA
7.2	Small town core: primary flow within an urban cluster of 2,500 to 9,999 (small UC)—Secondary flow 30% to 50% to a large UC
8.0	Small town high commuting: primary flow 30% or more to a small UC—No additional code
8.1	Small town high commuting: primary flow 30% or more to a small UC—Secondary flow 30% to 50% to a UA
8.2	Small town high commuting: primary flow 30% or more to a small UC—Secondary flow 30% to 50% to a large UC
9.0	Small town low commuting: primary flow 10% to 30% to a small UC—No additional code
10.0	Rural areas: primary flow to a tract outside a UA or UC—No additional code
10.1	Rural areas: primary flow to a tract outside a UA or UC—Secondary flow 30% to 50% to a UA
10.2	Rural areas: primary flow to a tract outside a UA or UC—Secondary flow 30% to 50% to a large UC
10.3	Rural areas: primary flow to a tract outside a UA or UC—Secondary flow 30% to 50% to a small UC

B.3.28 RX_ORIGIN_TYPE

[From CDM]

Table 45. RX_ORIGIN_TYPE Details

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

B.3.29 SETTING_TYPE

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

Table 46. SETTING_TYPE Details

Code	Definition
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]

B.3.30 SEX_TYPE

[From CDM]

Table 47. SEX_TYPE Details

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

B.3.31 SPECIALTY_TYPE

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

B.3.32 SPECIMEN_SOURCE_TYPE

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

B.3.33 UNIT_TYPE

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

B.3.34 VITAL_SOURCE_TYPE

[From CDM]

Table 48. VITAL_SOURCE_TYPE Details

Code	Definition
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

B.3.35 YES//NO

[From CDM]

Table 49. YES//NO Details

Code	Definition
Y	Yes
N	No
R	Refuse to answer

Code	Definition
NI	No information
UN	Unknown
OT	Other

Appendix C. RLDM Documentation

C.1 Overview

The RLDM introduces two ancillary tables that support privacy preserving record linkage shown in Figure 7.

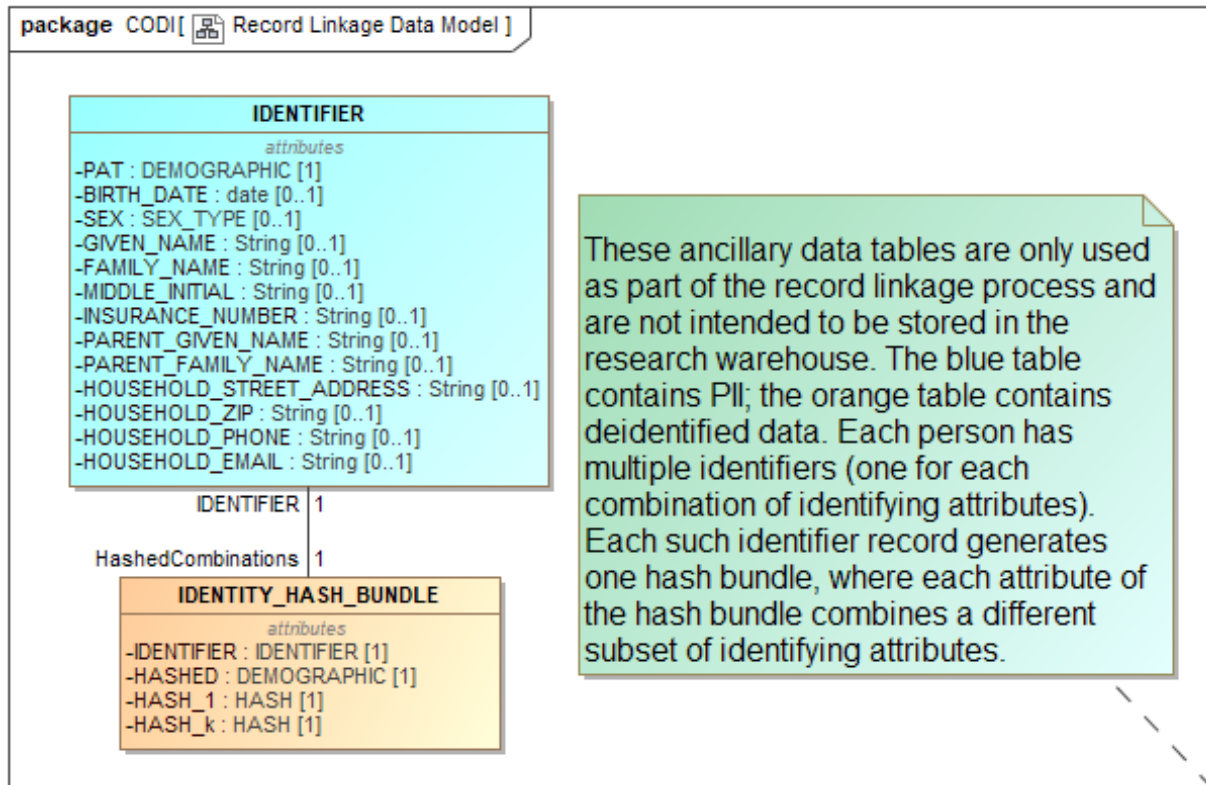


Figure 7. CODI Record Linkage Ancillary Tables

C.2 IDENTIFIER

The IDENTIFIER table contains one row for each unique combination of identifying information. The intention is that this table can be stored separately from research data because it has PII. The exact configuration will be determined experimentally. These experiments will determine which attributes to include and how those attributes will be normalized by each site (prior to hashing).

Table 50. IDENTIFIER Details

Attribute	Cardinality	Type	Definition
IDENTIFIERID	1	ID	
PATID	1	FK::DEMOGRAPHIC	A link back to the demographic table.
BIRTH_DATE	0..1	date	Date of birth.
SEX	0..1	SEX_TYPE	Sex assigned at birth.
GIVEN_NAME	0..1	String	A given name for the child.
FAMILY_NAME	0..1	String	A family name for the child.
MIDDLE_INITIAL	0..1	String	A middle initial for the child.
INSURANCE_NUMBER	0..1	String	An insurance number for the child as it appears on an insurance card.
PARENT_GIVEN_NAME	0..1	String	A given name for a parent of the child.
PARENT_FAMILY_NAME	0..1	String	A family name for a parent of the child.
HOUSEHOLD_STREET_ADDRESS	0..1	String	An address for the child, including number/name/unit (i.e., the information sometimes referred to as street line 1 and street line 2).
HOUSEHOLD_ZIP	0..1	String	A ZIP code for the child.
HashedCombinations	1	FK::IDENTITY_HASH_BUNDLE	A hashed version of this identifier record.
HOUSEHOLD_PHONE	0..1	String	A phone number for the child.

Attribute	Cardinality	Type	Definition
HOUSEHOLD_EMAIL	0..1	String	An email address for the child.

C.3 IDENTITY_HASH_BUNDLE

The IDENTITY_HASH_BUNDLE table contains one record for each record in the IDENTIFIER table. By storing the hash values into their own table, this table can appear in a data warehouse instance that eschews PII.

Table 51. IDENTITY_HASH_BUNDLE Details

Attribute	Cardinality	Type	Definition
IDENTIFIERID	1	FK::IDENTIFIER	A link back to the IDENTIFIER table.
HASHEDID	1	FK::DEMOGRAPHIC	A hashed version of the PATID.
HASH_1	1	HASH	A hash value for configuration 1 of identifier values.
HASH_k	1	HASH	A hash value for configuration k of identifier values.

Acronyms

BMI	Body Mass Index
CDC	Centers for Disease Control and Prevention
CDM	Common Data Model
CHORDS	Colorado Health Observation Regional Data Service
CODI	Childhood Obesity Data Initiative
DCC	Data Coordinating Center
EHR	Electronic Health Record
ETL	Extract–Transform–Load
FFRDC	Federally Funded Research and Development Center
IT	Information Technology
OMOP	Observational Medical Outcomes Partnership
PCORnet	Patient Centered Outcomes Research Network
PII	Personally Identifiable Information
PPRL	Privacy-Preserving Record Linkage
RDM	Research Data Model
RLDM	Record Linkage Data Model
VDW	Virtual Data Warehouse

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 3.1 Data Model Manual*, 2018. Denver, CO.

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 v5.1 Specification*, Sep. 12, 2019. Available:

https://pcornet.org/wp-content/uploads/2019/09/PCORnet-Common-Data-Model-v51-2019_09_12.pdf

V. Rogers, P. Hart, E. Motyka, E. Rines, J. Vine, D. Deatrack. *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

NOTICE

This document was produced for the U. S. Government under Contract Number HHSM-5000-2012-00008I, and is subject to Federal Acquisition Regulation Clause 52.227-14, Rights in Data-General.

No other use other than that granted to the U. S. Government, or to those acting on behalf of the U. S. Government under that Clause is authorized without the express written permission of The MITRE Corporation.

For further information, please contact The MITRE Corporation, Contracts Management Office, 7515 Colshire Drive, McLean, VA 22102-7539, (703) 983-6000.

© 2020 The MITRE Corporation.