

3. Table and Field Definitions and Descriptions

This chapter contains a complete reference list of tables and fields in the i2b2 CDM, followed by a detailed description of each of the tables.

3.1 Schema Reference

A spreadsheet describing all the tables and fields in the i2b2 CDM can be found at

<https://docs.google.com/spreadsheets/d/1ZR6X6JUHS-uC5Lz5rgBGi0vEJp99gGjDLLMH7Dnz0FLg>

It contains a row for each table and field. The columns "**Data Type: MSSQL**", "**Data Type: Oracle**", and "**Data Type: Postgres**" indicate the data type of the field in Microsoft SQL Server (MSSQL), Oracle, and Postgres, respectively. The "**Primary Key**" column indicates which fields are primary or foreign keys.

The "**Core**" column indicates the fields that are currently used by the i2b2 and tranSMART software and required to be available in any implementation of the i2b2 CDM. The "**Admin**" column indicates administrative fields that are used to indicate the source system for the data, when data were imported into the table, when the data were last updated, etc. These fields are not used by query or analysis features within the i2b2 or tranSMART software. However, some ETL tools for i2b2 and tranSMART use them. The "**Future**" column indicates fields that are intended to be used in future versions of i2b2 and tranSMART. Fields that are neither "Core", "Admin", nor "Future" are optional. They represent common types of data. For example, the "sex_cd" and "race_cd" fields in the PATIENT_DIMENSION table are often used to store the sex and race of the patient. However, these fields are not required, and the software does not specify what codes to use. The customizable ontology in i2b2 and tranSMART defines whether these fields are used, what codes are allowed, and what those codes mean.

Additional fields can be added to any of the tables in the i2b2 CDM. The ontology defines how they are used. Any field that is not a Core field can be removed. For large amounts of data (e.g., billions of observations), removing unused Core fields can save significant disk space. Additional dimension tables can also be created. Again, the ontology defines how these dimension tables link back to the OBSERVATION_FACT table.

The "**Values**" column indicates fields for which certain values are required. The "**Description**" column contains a brief description of the field. The "**History**" column indicates the first version of i2b2 that contained the column. Note that for backwards compatibility, no fields have ever been removed from i2b2. The "**Notes**" column contains additional information about certain fields.

3.2 Table and Field Descriptions

Each of the i2b2 CDM tables have a set of **administrative columns**, which are primarily used by ETL processes. A description of these fields are at

<https://community.i2b2.org/wiki/display/ServerSideDesign/General+Information>

Detailed descriptions of the **OBSERVATION_FACT** table is at

https://community.i2b2.org/wiki/display/ServerSideDesign/OBSERVATION_FACT+Table

Descriptions of the **five dimension tables** are at

https://community.i2b2.org/wiki/display/ServerSideDesign/PATIENT_DIMENSION+Table

https://community.i2b2.org/wiki/display/ServerSideDesign/VISIT_DIMENSION+Table

https://community.i2b2.org/wiki/display/ServerSideDesign/PROVIDER_DIMENSION+Table

https://community.i2b2.org/wiki/display/ServerSideDesign/CONCEPT_DIMENSION+Table

https://community.i2b2.org/wiki/display/ServerSideDesign/MODIFIER_DIMENSION+Table

The **relationships** between these dimension tables and the fact table are described at

<https://community.i2b2.org/wiki/display/ServerSideDesign/Joining+Columns>

The i2b2 software uses two optional mapping tables, **PATIENT_MAPPING** and **ENCOUNTER_MAPPING**, which map the i2b2 patient_num and encounter_num to local codes, such as medical record numbers or encounter billing IDs. These tables are not part of the core i2b2 CDM, but they are used by many sites. Descriptions of these tables are at

https://community.i2b2.org/wiki/display/ServerSideDesign/PATIENT_MAPPING+Table

https://community.i2b2.org/wiki/display/ServerSideDesign/ENCOUNTER_MAPPING+Table

3.3 Demo Data (Online)

i2b2 comes with a sample data set of 133 fake patients. This sample data is viewable as a set of Google Sheets at the following link:

<https://drive.google.com/drive/folders/11a6Yr-T9DSj0KsJ9XP4HQwgbwTfORRQu?usp=sharing>