ESP – MDPHNet Data Model

High Level Specification

Version 1.0

**Revision History**

| Version Number | Modification Date | By | Description of Changes |
| --- | --- | --- | --- |
| 0.3 | 13 Mar 2012 | R. Schaaf | Initial version |
| 1.0 | 09 Jul 2012\3 | B. Zambarano | Production release |
|  |  |  |  |
|  |  |  |  |

Contents

[1 Introduction 1](#_Toc319425233)

[2 Data Model Components 1](#_Toc319425234)

[3 Patient-Level Tables 1](#_Toc319425235)

[3.1 ESP\_DEMOGRAPHIC 2](#_Toc319425236)

[3.2 ESP\_ENCOUNTER 3](#_Toc319425237)

[3.3 ESP\_DIAGNOSIS 4](#_Toc319425238)

[3.4 ESP\_DISEASE 5](#_Toc319425239)

[4 Unique Value Tables 6](#_Toc319425240)

[5 Summary Tables 8](#_Toc319425241)

[5.1 ESP\_DIAGNOSIS\_ICD9\_3DIG 8](#_Toc319425242)

[5.2 ESP\_DIAGNOSIS\_ICD9\_4DIG 9](#_Toc319425243)

[5.3 ESP\_DIAGNOSIS\_ICD9\_5DIG 10](#_Toc319425244)

# Introduction

This design note proposes a data model for use by the MDPHNet application to query data stored in ESP. This data model is simpler than the full ESP data model and is motivated by a desire to:

* Provide a mapping layer between the full set of ESP tables and the MDPHNet application to allow the organization of ESP data to evolve over time without adversely impacting MDPHNet.
* Adapt the ESP data model to provide a set of database tables that are better suited to the needs of the MDPHNet query application.
* Improve database query performance relative to what would have been achievable running queries against the full set of ESP tables.

# Data Model Components

The simplified data model supplies:

* A set of patient-level tables. For the initial phase of the MDPHNet project, these will include demographic, encounter, diagnosis and detected disease data. More patient-level tables can be added in later project phases.
* A set of unique value tables (UVTs) that the application can leverage to efficiently gather lists of categorical (i.e., ”pick list”) information.
* A set of summary tables that provide counts of visits and patients stratified by center, age group, time period, encounter type and diagnosis.

# Patient-Level Tables

The ESP – MDPHNet data model provides the following tables.

Where relevant, the organization and format of information in these tables is patterned after the Mini-Sentinel Common Data Model, v2.1.

* **ESP\_DEMOGRAPHIC**  
  Contains one record per patient with the most recent information on birth date, sex and race.
* **ESP\_ENCOUNTER**  
  Contains one record per patient visit. This table includes provider, visit date, visit location, encounter type and age group information.
* **ESP\_DIAGNOSIS**

Contains one record per combination of patient, visit and diagnosis.

* **ESP\_DISEASE**

Contains one record per combination of patient and detected disease. The table includes detection criteria, review workflow status, reviewer notes and derived variables for 5 year, 10 year and Mini-Sentinel age groups.

## ESP\_DEMOGRAPHIC

This table contains the most recent patient demographic information on birth date, sex and race.  
  


PATID is used as a primary key constraint

There are seven btree indexes on the table, one for each field.

CENTERID, RACE and SEX are constrained by foreign key constraints against their respective UVT tables described below.

## ESP\_ENCOUNTER

This table contains patient visit information along with derived variables for 5 year, 10 year and Mini-Sentinel age groups.



ENCOUNTERID is used as a primary key constraint

There are fourteen btree indexes on the table, one for each field.

AGE\_GROUP\_5YR, AGE\_GROUP\_10YR, AGE\_GROUP\_MS, ENC\_TYPE, ENC\_YEAR, FACILITY\_CODE, and PROVIDER are constrained by foreign key constraints against their respective UVT tables described below.

PATID is constrained by foreign key constraint against ESP\_DEMOGRAPHIC.PATID.

## ESP\_DIAGNOSIS

This table contains patient diagnosis information including the 3-digit and 4-digit ICD9 codes corresponding to the diagnosis. For ease of querying, this table also incorporates information from the visit associated with each diagnosis.

N.B.: DX, DX\_CODE\_4DIG, DX\_CODE\_4DIG\_WITH\_DEC, DX\_CODE\_5DIG, and DX\_CODE\_5DIG\_WITH\_DEC may include codes with less than 4 or 5 digits, if that reflects the corresponding code in the medical record.

PATID, ENCOUNTERID, DX comprise the primary key constraint

There are nineteen btree indexes on the table, one for each field.

There are two btree varchar\_pattern\_ops indexes on the table, one for DX, another for DX\_CODE\_5DIG.

DX, DX\_CODE\_3DIG, DX\_CODE\_4DIG\_WITH\_DEC, and DX\_CODE\_5DIG\_WITH\_DEC are constrained by foreign key constraints against their respective UVT tables described below.

PATID is constrained by foreign key constraint against ESP\_DEMOGRAPHIC.PATID.

ENCOUNTERID is constrained by foreign key constraint against ESP\_ENCOUNTER.ENCOUNTERID.

## ESP\_DISEASE

This table contains one record per combination of patient and detected disease along with the detection criteria, review workflow status, reviewer notes and derived variables for 5 year, 10 year and Mini-Sentinel age groups.



PATID, CONDITION, DATE comprise the primary key constraint

There are ten btree indexes on the table, one for each field except NOTES.

CONDITION, CRITERIA and STATUS are constrained by foreign key constraints against their respective UVT tables described below.

PATID is constrained by foreign key constraint against ESP\_DEMOGRAPHIC.PATID.

# Unique Value Tables

The data model includes a set of unique value tables (UVTs) as a convenience for the application to use in populating “pick lists” of categorical information. The format of the UVT tables is generally:



In addition, the two UVTs for DX\_4DIG and DX\_5DIG have an additional field called “item\_code\_with\_dec”, which includes the ICD9 code with the decimal point.

UVT tables are provided for the follow items.



# Summary Tables

The data model includes summary tables that provide counts of visits and patients stratified by center, age group, time period, encounter type and diagnosis.

In addition, there is a table summarizing Influenza-like Illness (ILI).

## ESP\_DIAGNOSIS\_ICD9\_3DIG

The ESP\_DIAGNOSIS\_ICD9\_3DIG table is analogous to the “3-Digit ICD-9 Diagnosis Summary Table Structure” defined as part of the Mini-Sentinel Common Data Model (v2.1).

This table provides a count of unique individuals stratified by age group, sex, year, and care setting who had at least one encounter recorded with the diagnosis of interest based on the first 3 digits of the diagnosis code. Three different sets of age groupings are represented in the summary table. The different age groupings are distinguished by “age\_group\_type”.



There are seven btree indexes on the table, one for each field except MEMBERS, EVENTS and DX\_NAME.

## ESP\_DIAGNOSIS\_ICD9\_4DIG

The ESP\_DIAGNOSIS\_ICD9\_4DIG table is analogous to the “4-Digit ICD-9 Diagnosis Summary Table Structure” defined as part of the Mini-Sentinel Common Data Model (v2.1).

This table provides a count of unique individuals stratified by age group, sex, year, and care setting who had at least one encounter recorded with the diagnosis of interest based on the first 4 digits of the diagnosis code. Three different sets of age groupings are represented in the summary table. The different age groupings are distinguished by “age\_group\_type”.  
  
Note: The summary table may also include 3 digit codes for those diagnoses that were entered using a 3-digit code.



There are seven btree indexes on the table, one for each field except MEMBERS, EVENTS and DX\_NAME.

## ESP\_DIAGNOSIS\_ICD9\_5DIG

The ESP\_DIAGNOSIS\_ICD9\_5DIG table is analogous to the “5-Digit ICD-9 Diagnosis Summary Table Structure” defined as part of the Mini-Sentinel Common Data Model (v2.1).

This table provides a count of unique individuals stratified by age group, sex, year, and care setting who had at least one encounter recorded with the diagnosis of interest based on the first 5 digits of the diagnosis code. Three different sets of age groupings are represented in the summary table. The different age groupings are distinguished by “age\_group\_type”.

Note: The summary table may also include 3 or 4 digit codes for those diagnoses that were entered using a 3 or 4-digit code.



There are seven btree indexes on the table, one for each field except MEMBERS, EVENTS and DX\_NAME.

## ILI\_SUMMARY

The ILI\_SUMMARY table is used to generate a set of standard reports for ILI syndrome surveillance. A row represents a weekly summary for a given age group at a given healthcare site.

|  |  |  |
| --- | --- | --- |
| **Column** | **Data Type** | **Description** |
| age\_group | text | Age groups in years 0-4, 5-24, 25-49, 50-64, 65+ |
| period\_end | date | For each week, date as of Saturday |
| week | character varying(5) | MMWR Week value. (See <http://wwwn.cdc.gov/nndss/document/MMWR_Week_overview.pdf>) |
| zip5 | character varying(5) | Five character zip code for healthcare site where diagnosis was made. |
| center | character varying(100) | Name of healthcare site where diagnosis was made |
| cdc\_site\_id | character varying(50) | Site ID as provided by CDC for reporting purposes |
| ili\_counts | integer | Count of ILI cases detected |
| tot\_counts | integer | Count of total patient encounters |