

---

# National Patient Information Reporting System: National Data Warehouse

---

## **NDW General Data Mart**

## **Getting Started Guide**

Version 3.0

April 2009



Department of Health and  
Human Services

Indian Health Service

Office of Information  
Technology (OIT)

---

# Contents

<b>Version Control.....</b>	<b>iii</b>
<b>Introduction.....</b>	<b>1</b>
About this Guide.....	1
<b>Accessing the General Data Mart.....</b>	<b>2</b>
Connection Configuration.....	2
Contact Information.....	2
<b>Access Privileges.....</b>	<b>3</b>
Database Access.....	3
Area Access .....	4
<b>General Data Mart Organization.....</b>	<b>4</b>
Patient Registration Data .....	5
Non-PII Registration Tables .....	6
Encounter Data.....	6
Reference Data.....	7
Administrative Data .....	7
<b>Data Retrieval – Best Practices.....</b>	<b>8</b>
Views .....	8
<b>Appendix A – New Table Structure .....</b>	<b>9</b>
New Data Column.....	9
Data Model.....	9

## Version Control

Version	Date	Notes
1.0	September 2005	Initial Draft
2.0	October 2005	Initial Release
2.1	November 2005	Changes based on customer feedback.
2.2	December 2007	Sub mart of Hollywood, add REF info
2.3	February 2008	Annual Review, removed Appendices (A-E) as information is available at the NDW web site at the following link:  <a href="http://www.ihs.gov/CIO/scb/metadata/">http://www.ihs.gov/CIO/scb/metadata/</a>  FY08 Bridge Contract deliverable (1.7.6); COTR accepted April 11, 2008
2.4	August 2008	Update for export_info and integrity ID additions by area access
3.0 Draft	March 2009	Add/Remove: <ul style="list-style-type: none"> <li>• schemas for the Restricted Personally Identifiable Information (PII) Data Requirements</li> <li>• data element that allows control of Area security</li> <li>• schemas for each version of past userpop reports run from NDW data</li> <li>• remove ssn from Pat_reg, userpop and enctrss tables</li> <li>• add new tables; pat_reg_ssn, userpop_ssn and enctrss_ssn</li> </ul> <p>COTR approval April 3, 2009</p>

## Introduction

The National Data Warehouse (NDW) **General Data Mart (GDM)** contains a sub-set of the NDW registration and encounter data. The contents of this data mart include:

- All current registration data contained in the data warehouse
- Full registration data which **includes** Personally Identifiable Information (PII)
- Non-PII registration data, which **excludes** Personally Identifiable Information (PII)  
Historic and current user population tables
- All current encounters contained in the data warehouse
- All reference tables
- Meta data, which contains information about the tables and columns included in the database
- Administrative data, including a timestamp for Mart refreshes and also information pertaining to file exports

The NDW General Data Mart is refreshed per the *HOLLYWD Database and Associated Data Marts Service Level Agreement, Version 2.0*. The General Data Mart also includes the Legacy Data Mart. Data in tables associated with the Legacy Data Mart is not refreshed.

## About this Guide

The *NDW General Data Mart Getting Started Guide* was developed to give experienced users a fundamental understanding of the General Data Mart. It is written specifically to address the needs of those users with more advanced statistical and analytical skills.

This guide does not provide complete documentation of the data mart, database use or concepts, nor is it a programming language tutorial.

## Accessing the General Data Mart

System access is enterprise compliant, allowing various environments to access the database, including ODBC, JDBC, OLE, and CLI. Software applications that can access databases using one of these standard protocols can connect to the General Data Mart. NPIRS will assist in obtaining a connection to the GDM but is not responsible for access programs.

## Connection Configuration

**Protocol:** TCP/IP  
**IP Address:** 198.45.1.8  
**Host Name:** bilbo.hqw.ihs.gov  
**Database Name:** HOLLYWD  
**Remote Instance Name:** db2hol

## Contact Information

To request access to the IHS General Data Mart, call or email

Stanley P. Griffith, MD, FAAFP  
NPIRS Program Manager, IHS National Data Warehouse Project  
Indian Health Service, Office of Information Technology  
**Phone:** 1-505-248-4144  
**E-mail:** [stanley.griffith@ihs.gov](mailto:stanley.griffith@ihs.gov)

Or send an email to: [NPIRSHD@ihs.gov](mailto:NPIRSHD@ihs.gov)

To promote a collaborative environment between the NDW staff and the users of the General Data Mart, the General Data Mart Users Group meets on a regular basis.

## Access Privileges

### Database Access

Only *authorized* users are allowed access to the General Data Mart, and will be assigned one of the following current security levels:

- National Level 1** access allows a user to view all data for all Areas.
- National Level 2** access allows a user to view partially cleansed data for all Areas.
- Area Level 1** access allows a user to view all data within their specified Area.
- Area Level 2** access allows a user to view partially cleansed data within their specified Area.

Additional security levels may be assigned or created in the future.

Partially cleansed data refers to tables with the PII removed. Tables may be identified by their schemas. The REG\_NP schema relates to data that is not generally PII sensitive (partially cleansed). The ENCTR schema generally does not contain PII data.

The table structure is located in the document titled *GDM\_Schemas\_TablesViewsNicknames* located at <http://www.ndw.ihs.gov/what-if-I-have-other-questions.asp>.

Below is a graphical depiction of the different levels of access.

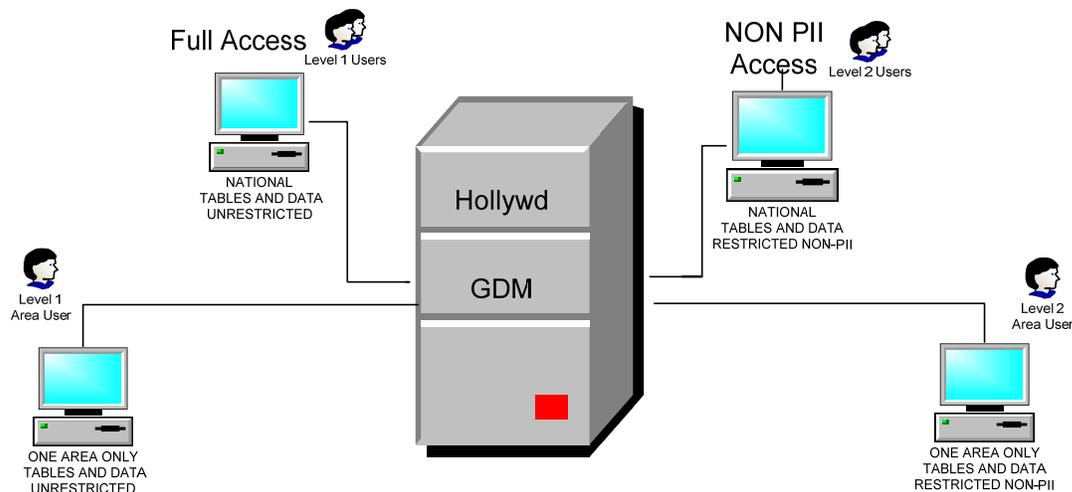


Figure 1. Illustration of the different levels of national and area access

In all cases, authorized users have **Read-Only** privileges.

## Area Access

Area Access is controlled by the Security Administrator and based on the user's ID. Area users may access data from the ENCTR, REG or REG\_NP schemas and will only be able to see data relevant to their security access.

## General Data Mart Organization

The General Data Mart is organized into the following “groups” that separate different kinds of data into their own spaces.

- REG contains the registration tables created to store information obtained from patient registration records received by the National Data Warehouse.
- ENCTR contains the encounter tables created to store information about patient encounter records received by the National Data Warehouse.
- REF contains both the Standard Code Book tables and the IHS code sets used by the National Data Warehouse (NDW). IHS Standard Code Book tables are available on our web site: <http://www.ihs.gov/CIO/scb/>. For a complete list of the Reference tables, see the *GDM Schemas and Tables/Views/Nicknames* document.
- META tables contain basic information on the structure and definition of columns and tables used in the mart. This information may be accessed via the Meta Data web site: <http://www.ihs.gov/CIO/scb/metadata/>.
- SYSIBM (SYSCOLUMNS, SYSTABLES) also contains meta data about the mart.
- ADMIN contains two tables:
  - LAST\_ETL – contains the dates of the extracts to refresh the mart.
  - EXPORT\_INFO – contains information about file exports.
- LEGACY tables contain static data from the legacy NPIRS system. This contains data collected prior to the startup of the NDW database. For additional information concerning the legacy tables and their usage, see the *Legacy Data Mart Getting Started Guide* Version 1.0. This data is historical data (Pre-NDW) and is not refreshed. For a listing and description of these tables, as well as models with physical and logical names, go to the IHS NDW Meta Data web site:

<http://www.ihs.gov/CIO/scb/metadata/>

All the tables and columns in the GDM can be viewed in the document, *GDM Schemas and Tables/Views/Nicknames* located on our website at: <http://www.ndw.ihs.gov/what-if-I-have-other-questions.asp>.

## Patient Registration Data

The **Patient Registration** tables contain information about the *individual* patients registered in the various IHS and tribal facilities.

---

**Note:** A patient may be registered at more than one facility, or more than once at the same facility, and therefore may have multiple registration records.  
**No un-duplication is applied.** The REG.USERPOP AND REG\_NP.USERPOP tables contain one data element called integrity id. This column contains a unique identifier that can be used to un-duplicate registrations belonging to the same patient within a given version number of USERPOP data. Multiple version numbers are contained within the table. A version number is a numeric code that represents a snapshot of registration data for a specific userpop report. An explanation of each version number is located in the USERPOP\_LOG table.

---

Within the REG schema, the PAT\_REG and DEMOGR tables are the tables that contain basic registration information about the patient. Typically, PAT\_REG holds more static patient information, such as Birth data, Gender, and SSN. DEMOGR holds changeable patient information, such as Name, Address, and Community of Residence.

The CHART, ALIAS, and INSUR\_ELIG tables contain additional information about the patient, if applicable, and there may be multiple records for that patient. Records in these tables utilize a sequence number field (SEQ\_NBR). The sequence number represents the order in which the data was received by the NDW.

The primary key used by the Patient Registration tables is **REG\_ID**. A Registration Identifier (REG\_ID) is the primary data element within all Registration tables that uniquely identifies each registration record in Registration tables within REG and REG\_NP schemas. This key is the unique identifier for the Patient Registration record, and is used to link all of the patient's registration-related data. The REG\_ID is generated by the NDW to ensure uniqueness of the registration within the database. The REG\_ID field is found in each of the Registration tables and may be used to retrieve registration information for the same registrant from multiple tables.

---

**Note:** Some encounter records are received without registration information. These records will have REG\_ID = 0 and cannot be linked to registration records.

---

Encounter information (discussed in the next section) can be linked to individual patient registration records for reporting purposes via the REG\_ID (when a registration record exists). The REG\_ID is the typical method of linking encounters and registration data, but other methods may be used, depending on the intent of the query.

## Non-PII Registration Tables

The REG\_NP schema has the same tables (except for the ALIAS table) as the REG schema; however, the PII has been removed. These tables accommodate security requirements for the General Data Mart users. Information that can directly identify a patient has been removed from these tables. Some PII data, such as birth dates and death dates, have been retained in order to assist the user in preparing reports.

The new tables and schemas are listed in Appendix A.

## Encounter Data

The ENCTRSS table contains the basic information about encounters. Other encounter tables contain more specialized information, such as dental and contract data. For example, the ADA table contains American Dental Association codes related to a dental visit. Explanations of these tables and what they contain can be found on our web site at: <http://www.ihs.gov/CIO/scb/metadata/>

Encounter tables may hold more than one row of data per encounter. The data is stored as received; **no un-duplication is applied**. For reporting purposes, a duplicate flag variable (DUP\_FG) is available for un-duplication of encounters.

For a complete listing and description of the Encounter tables, as well as models with physical and logical names, go to the IHS NDW Meta Data web site: <http://www.ihs.gov/CIO/scb/metadata/>

A NDW Encounter Identifier (ENCTRSS\_ID) is a key data element that uniquely identifies each encounter record in Encounter tables within ENCTR schema. The ENCTRSS\_ID is generated by the NDW to ensure uniqueness of the encounter within the database. The ENCTRSS\_ID field is found in each of the Encounter tables and may be used to retrieve encounter information stored in multiple tables.

A sequence number (SEQ\_NBR) is stored for multiple records of the same encounter. This number is the order in which NDW receives the data.

The main Encounter table (ENCTRSS) contains REG\_ID field which indicates the NDW registration identifier for the encounter data.

Encounter records are usually linked to a patient registration in Registration schemas by registration ids stored in the main encounter table (ENCTRSS). Some encounter records will not have registration IDs due to timing issues in receiving encounter and registration data by NDW. The registration ids (REG\_ID) for those records will default to 0. Prior to each GDM refresh there will be an attempt to link these unlinked encounters to a newly received registration.

## Reference Data

The reference tables contain descriptions and other pertinent data concerning the codes utilized in the NDW. The reference tables are in the REF schema. The code values are generally listed as a column name ending with \_CD. For example, the FACILITY table contains the facility codes (FAC\_CD) and names (FAC\_NM) for facilities. In addition, some tables contain a flag indicating whether data in a column is workload reportable.

The reference tables reflect codes that are used in the RPMS and NDW system. They may not be the most current list since the NDW does not receive regular updates from RPMS for most tables. Check with your local system administrator for a complete listing.

For a complete list of the tables with descriptions, see the *NDW Reference Tables* document at <http://www.ndw.ihs.gov/what-if-I-have-other-questions.asp>.

## Administrative Data

Administrative tables are referenced by the ADMIN schema. The LAST\_ETL table contains the dates of the extracts that were performed to refresh the mart and is used for informational purposes only. The EXPORT\_INFO table contains data about the file exports, including EXPORT\_ID and various dates.

## Data Retrieval – Best Practices

### Views

Database views may be used to organize the data in alternative ways, depending on user needs. When doing so, keep in mind that there are two different kinds of views, and this affects the security for the view (who is authorized to view the data contained and who may alter the views).

- 1) **Private** – If the view is created using the creator's userid as the schema (this is the default), the view and data will be made accessible only to the person creating the view.
- 2) **Public** – If the view is created using ENCTR, REG,REG\_NP, or NATL, the view and data will be accessible to users having the appropriate access.

Views should be used only when they are expected to remain persistent (i.e., used over and over, to resolve very complex SQL code), or if they are expected to be used in a stored procedure. In general, a user is expected to drop (delete) views that are no longer needed.

## Appendix A – New Table Structure

The tables below have been added to accommodate the security requirements for the General Data Mart. A schema for Registration tables with Non-PII (REG\_NP) has been created for tables that have been partially cleansed of Personally Identifiable Information (PII). The REG\_NP schema identifies registration tables for security Level 2 users.

The REG\_VERx schemas are specifically designed for each version of the userpop reports. These tables are only available for National Level users and were added to expedite data pulls.

The following tables have been added :

- REG\_NP.PAT\_REG
- REG\_NP.DEMOGR
- REG\_NP.INSUR\_ELIG
- REG\_NP.CHART
- REG\_NP.USERPOP
- REG.PAT\_REG\_SSN\*
- ENCTR.ENCTRSS\_SSN\*
- REG.USERPOP\_SSN\*
- REG\_VERx.USERPOP

*\* The REG.PAT\_REG\_SSN, REG.USERPOP\_SSN and the ENCTR.ENCTRSS\_SSN tables will have limited user access based on the security approval obtained by the user.*

### New Data Column

A new data column, security\_tag, has been added to all tables. This column can not be used for reporting purposes. It is a column that allows the Security Administrator to control appropriate security access to the tables.

### Data Model

All the tables and columns in the GDM can be viewed in the document, *GDM Schemas and Tables/Views/Nicknames* located on the website at: <http://www.ndw.ihs.gov/what-if-I-have-other-questions.asp>