
National Data Warehouse

Area Preparation Guide

*Provides Requirements and Guidelines to Customers Preparing
for the Initial Data Load to the National Data Warehouse*

Version 2.0

August 2004



Department of Health and
Human Services

Indian Health Service

Information Technology
Support Center
(ITSC)

Contents

Introduction.....	1
Getting Help	1
Getting Started.....	1
Software Requirements.....	2
The Initial NDW Extracts	3
Initial Load Process Summary.....	4
Runtime and Hard Disk Space Statistics	5
Area NDW Preparation Checklist.....	7



Introduction

The purpose of this guide is to outline for the applicable Facility/Area staff the steps necessary to prepare for sending data that will be placed in the IHS National Data Warehouse (NDW) repository. The information in this guide is based on all data processed so far and contains all known information that is necessary to prepare for a successful Initial Data Load to the NDW. As the process evolves, this guide will be revised, and new versions will be posted on the NDW website.

Please note that during the Initial Data Load phase, facilities must FTP the files manually. However, when the RPMS data export process goes into full production mode, the process will run automatically.

Getting Help

For additional information, see the National Data Warehouse website:

<http://www.ndw.ihs.gov/>

Please email any questions or suggestions regarding this guide to:

itscdatawarehouse@ihs.gov

Getting Started

Make sure the required RPMS software is installed

Begin surveying all Area facilities for what RPMS software updates are needed and apply those updates. For details see the “Software Requirements” section on page 2.

Provide Information NDW Needs

Approximately **three** weeks before your scheduled start date, we will request that the Area ISC provide to us the following information.

- **Contact list** We need a **contact name**, **email address**, and **phone number** for each facility that will be sending data to the NDW.

The Contact list allows us to communicate to the facility information related to the process, such as an acknowledgement that the files have been received or a phone call to the contact to discuss support issues or file rejections.



-
- **Facility list** We need the **name** and **ASUFAC** of each facility that will be sending data to NDW.
 - **Project lead** If we do not have the correct Area Project Lead, please email the Initial Data Load project lead for your area.

Review the proposed schedule and make changes

The NDW provides the final schedule of when each facility is to start sending its exports. Your input into creating this schedule is very important.

The facility list that NDW creates is list based on the total number-of-visits count per facility, where the facility with the largest count is first.

When we receive from you the requested Area Facility List, we will compare lists and send to the Area a proposed schedule of when their facilities are to start their initial exports.

You should review this schedule and change the order based on the level of difficulty that the facility may encounter, where those facilities with the most hurdles would start earlier in the cycle; or any other scheduling conflicts you may foresee.

If the number of visits per month and the total number of registration records is known by facility, you can derive how long the processing might take and how much space is required in MB to perform the initial exports. To compute the needs of your facilities, see the “Runtime and Hard Disk Space Statistics” section on page 5.

Software Requirements

Prior to the initial load, install the following patches. Installing these patches is *critical*, especially the Patient Registration patches

Important: Please apply patches *in order* within a group. For example, apply AVA patch 17 before patch 18.

The RPMS minimum software requirements are:

1. GIS Version 3.01
2. GIS Version 3.01 patch 2
3. Fileman version 22
4. Kernel 8.0 or later
5. IHS Patient Dictionaries (AUPN) v 99.1 through patch 13 (when patch 14 is released that will be a requirement)



6. IHS VA Support Files (AVA) v 93.2 through patch 18.

NOTE If you are a PIMS user, AVA patch 18 must be followed immediately by an install of PIMS.

7. PCC Management Reports v3.0 through patch 15

8. RPMS Registration v7.0 through patch 2

The Initial NDW Extracts

The initial load, or “back load,” to the NDW consists of:

- A one time full registration extract
- Visit extract(s) containing visits from October 1, 2000 to the present

Considerations

The Registration and Visits extracts are created using HL7 messages. The HL7 messages will be output to a host file for export to the NDW computer.

NOTE The HL7 messages are fairly large and will use up disk space during the initial load process.

The Visits extract takes up space in the Cache environment (^INTHU global and temp globals) and creates an output file that you must FTP to the NDW computer. The space needed is temporary because there is a purger, which will purge all HL7 messages from the INTHU global once they are 3 days old.

The facility will need to manage the archiving of the files that are created.

Sending the Initial Extracts to NDW

The NDW will supply to the Area ISC the address and login instructions needed to FTP the export to NDW.

During the Initial Load phase of the RPMS data export process, each facility must FTP the generated files manually. Please send both the audit and extract files that are generated during the initial load.

After you receive an acknowledgement from NDW that your files were received, you may safely delete the corresponding local audit and extract files, because the files you sent are now stored in the NDW.



Initial Load Process Summary

NOTE The IHS/RPMS NDW software used for the Initial Load process will be available three weeks prior to your scheduled start date.

The following steps summarize the basic process for running initial loads (vs. incremental).

1. Install the IHS/RPMS NDW software package using kids build bdw_0100.k.
2. Install GIS patch 12 using kids build gis_0301.12k, gis_0301.12n, and gis_0301.12g.
3. Assign BDW security keys to the users responsible for NDW initial loads and for incremental NDW exports. The keys are:

Key	Description
BDWZMEN	Unlocks the main menu
BDWZ RESET	Enables re-setting a log entry
BDWZ REG EXPORT	Enables the user to do a full registration export

4. Run the full Registration initial load using the ARG Data Warehouse Full Registration Export option. This option is on RDWMENU, using the patch BLM-ARG.

Two files will be created and placed into the defined export directory of the site:

- BDWDWPXasufac.julian date (first file name)
- asufacdatetime.BDW (second file name)

An example of these files from SX:

BDWDWPX000111.123
0001113040530112045.BDW

5. Perform the Visit initial load.

The visit extract is created using the BLV menu option. Based on the size of the site and the available disk space, each facility must determine how many visits to process with each run. For information on determining how many visits to process, see the “Runtime and Hard Disk Space Statistics” section on page 5.

We are aware that there is a lag between the time when a service was actually provided and when the related data is entered into the local system. We want to assure you that any encounter data that has not been entered up to the date of the extract will be included in subsequent incremental loads.

Therefore, it is important and necessary to run the last encounter export through (and including) the designated end date, which is based on the registration extract.



To find out your registration run date, so you can finish the Initial Load process for encounters:

Run the process and at the “what ending date” prompt, type **T** and press the Enter key.

For example, if the date of the registration extract is 09/23/2004, the on-screen response would be: “Response must not follow 09/23/2004.”

The date that appears is the date to use as the ending date for the final encounters run of the Visit initial load.

If you need more information or detailed instructions on running the processes, see the *NDW Data Load Processing Technical User Guide*

Runtime and Hard Disk Space Statistics

The following statistics were gathered using all data that was available at the time of this version of the *Area Preparation Guide*. Our data assumes the exports were executed after hours and no other processes were running. Other factors that could affect the results include the type of hardware, processor, and throughput speeds, the amount of memory or HD space, and the number of applications running on the same server or at the same time.

1) Registration – Initial Loads

This process is run one time, creating an audit file and an extract file in the same process. To derive a runtime and to determine the amount of MB space needed, you need to know approximately how many registration records you have at your facility.

a) Runtime

Where Y = the number of these records:

$(Y * 4 \div 10,000)$ = the total runtime in minutes to create both files.

Example If you have 200,000 registration records,

$(200,000 * 4 \div 10,000) = 80$ minutes.

80 minutes divided by 60 minutes = 1 hours and 20 minutes

b) Space in MB

Where Y = the number of these records:

$(Y * 14.5 \div 10,000)$ = Total space in MB for both files in total.

Example If you have 200,000 registration records,

$(200,000 * 14.5 \div 10,000) = 290$ MB (megabytes of disk space).



2) Encounters (Visits) – Initial Loads

This process creates an audit file and an extract file in the same procedure. Because there can be many more records than the registrations, this process can be run multiple times by date increments.

To keep the processing time manageable, you can run the process by a date range. To derive a runtime and to determine the amount of MB space needed, you need to know approximately how many visits you have at your facility each month on average.

a) Runtime

Where Y = the number of these records each Month * 6 (for six months):

$(Y * 10.6 \div 10,000)$ = Total space in MB of both files in total.

Example If you have 7,000 visits each month,
 $(7,000 * 6) = 42,000$
 $(42,000 * 10.6 \div 10,000) = 44.5$ MB (megabytes of disk space).

b) Space in MB

Where Y = the average number per month of these records * 6 (for six months):

$(Y * 14 \div 10,000)$ = the total runtime in minutes to create both files.

Example If you have, on average, 7,000 visits each month,
 $(7,000 * 6) = 42,000$
 $(42,000 * 14 \div 10,000) = 59$ minutes (rounded)

Area NDW Preparation Checklist

- Reviewed RPMS software update status at each facility, and installed patches as needed.
- Sent list of Area facilities/ASUFACs to NDW.
- Sent facility contact information to NDW.
- Performed calculation by facility for runtime and space constraints, if needed.
- Received preliminary Initial Data Load schedule from NDW staff.
- Reviewed and proposed changes to the Initial Data Load schedule.
- Received final Initial Data Load schedule from NDW staff.
- Received and installed the IHS/RPMS NDW software from NDW staff.
- Received FTP instructions from NDW staff.

