

Introduction

This application has been developed to assist VAMC facility representatives in identifying and merging duplicate records found in *VISTA* files. Record pairs are identified as potential duplicates through a search of the database. These potential duplicates are then validated through a review process to verify that they are duplicates, then merged. This application is intended to provide a reliable approach to correctly identify and merge duplicate records.

In order to competently operate the application you must be familiar with the operations of the *VISTA* computer system in general. If you do not use the system on a regular basis, it is recommended that you review the **User's Guide to Computing**.

A detailed understanding of VA FileMan is not required to successfully use this application. However, reviewing the **VA FileMan User's Manual** provides you with a good background for how the system works.

On-line help is provided at all prompts by typing one or two question marks.

Product Description

Patient Merge provides an automated method to eliminate duplicate patient records within the *VISTA* database. It is an operational implementation of the Duplicate Resolution Utilities, which were released to the field with Kernel Toolkit.

The overall process consists of three major subject areas: the search for potential duplicate record pairs, review, verification, and approval of those pairs, and the merge process.

The search and identification of potential duplicate records performs comparisons on key fields in the PATIENT file. By selecting a Basic search option you begin a search on a file for the first time. Routines are activated containing preset specifications for duplicate tests. These tests are identified in the DUPLICATE RESOLUTION file (#15.1). Each test uses its corresponding field and file numbers to assist in identifying potential duplicate records. Those fields are Name, Social Security Number, Sex, Date of Birth, Date of Death, Last Separation Date (Last Discharge Date), Mother's Maiden Name, and Claim Number. A value is ascribed to the field based upon similarity in the data present. Then all field values are tabulated to form a single value, which is compared against the Threshold Percentage (a site parameter). When record pair scores evaluate above this percentage, they are considered to be potential duplicates, entered into the

DUPLICATE RECORD file (#15), and made available for further processing. IRM personnel authorized to possess the XDRMRG security key generally initiate this task.

The search process is a CPU intensive job that has been known to exceed 100 hours depending upon factors such as the number of entries in the PATIENT file and the hardware involved. Utilities are provided to allow for pausing and restarting the job.

Once the search has been initiated and has found a potential duplicate pair, the process of verifying record pairs begins. The review and verification process may begin while the search is running.

The review and verification process includes two levels of review. The primary reviewer, initially seen as an MAS responsibility, performs a review of patient demographic information. The primary reviewer initially determines if the pair represents a duplicate record. If so, the primary reviewer selects the merge direction. If data from ancillary services is present, notification (via MailMan message or alert – or both) is sent to those designated as ancillary reviewers. A site may determine reviewers based upon their business practices. Reviewers determine whether the record pair is a duplicate, not a duplicate (so that subsequent processing need not occur), or that they are unable to determine the status. Where appropriate, reviewers may mark data to be overwritten. Those record pairs that are determined to be verified duplicates are marked as such and are then available for approving to be merged.

The intent of the approval step is to ensure that a conscious decision will be made in taking verified duplicate record pairs and making them available for a merge process. All verified record pairs, or selected pairs, can be approved. The approval step follows a site defined waiting period. Reviewers are responsible for approving verified duplicates.

The merge process is available for initiation by IRM personnel. All approved record pairs are included in a merge process when scheduled. The merge process is a lengthy process that is recommended for off-peak hours. Utilities are available for pausing and restarting the merge process. The merge process merges verified duplicate records in the following order: first, files that require special handling, then the primary file, then the resolution of pointers. The resolution of pointers for the primary file or any of those involving special processing involve three phases. The first two phases permit identification of entries requiring modification based on their IENs (DINUMed) or by cross-references and are fairly rapid. The third phase involves all other pointers and can be lengthy. Several special processing routines have been written to handle those database entries that point to the PATIENT file in an unusual manner. Entries for each special processing routine have been made in the PACKAGE file (#9.4) multiple, AFFECTS RECORD MERGE (Field #20). A

stub record is maintained in order to disallow reuse of PATIENT file internal entry numbers.

Concurrent with the merge, entries are made in a new global for each record making up the pair. The entries are intended to provide a "before-merge" image. However, please note that the merge is a non-reversible process. Once the pair of records is merged, there is no automated way of undoing the process.

The application has been written to support multiple parallel jobs (threads - as specified by the site) during the merge process. However, decreased overall processing time is exchanged for increased system utilization.

Advisory Note: The merge process is a background job. Be aware that it should not be running when changes are being made to Data Dictionaries or when data conversions are taking place.

Introduction

Orientation

How to Use this Manual

This manual is intended for use in conjunction with the Duplicate Record Merge: Patient Merge application. Items included in the release of Patient Merge, such as routines and files, are briefly described for quick reference.

To gain a comprehensive understanding of this application, read the *Duplicate Record Merge: Patient Merge User Manual* in conjunction with using the application.

This manual uses several methods to highlight different aspects of the material. "Snapshots" of computer dialogue (or other on-line displays) are shown in a non-proportional font and enclosed within a box. User responses to on-line prompts are highlighted in boldface. Boldface is also used to highlight a descriptive word or sentence. The Return or Enter key is illustrated by the symbol <RET> when displayed in computer dialogue and is included in examples only when it may be unclear to the reader that such a keystroke must be entered. The following example indicates that you should type two question marks followed by pressing the Return key when prompted to select an option:

```
Select Primary Menu option: ??
```

Figure 1 - How to access online help

M code, variable names, acronyms, or the formal name of options, actual field names, file names, and security keys (e.g., XDR, XDRMGR, and DG ELIGIBILITY) are represented with all uppercase letters.

Who Should Read this Manual?

This manual was written with many job functions in mind.

Since each site will determine who will control the patient merge process, what considerations are followed, how the merge will be accomplished, and when the merge should be started/stopped, everyone involved with the merge should read this manual.

Orientation

Implementation and Maintenance

Duplicate Record Merge: Patient Merge is a Kernel Installation and Distribution System (KIDS) software release.

The Duplicate Record Merge: Patient Merge application requires a standard **VISTA** operating environment in order to function correctly. Check your **VISTA** environment for packages and versions installed.

Minimum application requirements are:

- VA FileMan V. 21.0
- Kernel V. 8.0
- Kernel Toolkit V. 7.3
- NDBI V. 1.0

VISTA patch installations needed to install the Duplicate Record Merge: Patient Merge application are:

Package	Designation
Immunology Case Registry	IMR*2.1*2
Inpatient Medications	PSJ*4.5*43
Integrated Billing	IB*2*94
Integrated Billing	IB*2*97
Lab Service	LR*5.2*205
Registration	DG*5.3*156
VA FileMan	DI*21*41
VA FileMan	DI*21*42

VISTA packages needed to fully utilize the capabilities of the Duplicate Record Merge: Patient Merge application are:

- Patient Information Management System (PIMS) V. 5.3
- Health Summary V. 2.5

The following **VISTA** packages have files that require special processing during the merge process. Additionally, a routine for National Database Integration is also included for special processing during the merge process.

- Lab Service V. 5.2
- Immunology Case Registry V. 2.0
- Integrated Billing V. 2.0
- National Database Integration (Consolidation Site activity)

Records in these files contain fields affected by the merging of the PATIENT file (#2), but are not identified as pointer fields. Routines for processing these files are sent with this application since the PATIENT file commonly points to these files. Your site can create additional routines for processing other files pointed to the PATIENT file. Any time an additional routine is created, it is necessary that an entry be made in the PACKAGE file (#9.4) in the AFFECTS RECORD MERGE subfile.

The search process for potential duplicate records is a CPU intensive job that has been known to exceed 100 hours depending upon factors such as the number of entries in the PATIENT file and the hardware involved. Utilities are provided to allow for pausing and restarting the job.

The merge process is available for initiation by IRM personnel. All approved record pairs are included in a merge process when scheduled. The merge process is a lengthy process that is recommended for off-peak hours. Utilities are available for pausing and restarting the merge process. The merge process merges verified duplicate records in the following order: first, files that require special handling then the primary file, then the resolution of pointers. The resolution of pointers for the primary file or any of those involving special processing involve three phases. The first two phases permit identification of entries requiring modification based on their IENs (DINUMed) or by cross-references and are fairly rapid. The third phase involves all other pointers and can be lengthy. Several special processing routines have been written to handle those database entries that point to the PATIENT file in an unusual manner. Entries for each special processing routine have been made in the PACKAGE file (#9.4) multiple, AFFECTS RECORD MERGE field (#20). A stub record is maintained in order to disallow reuse of PATIENT file internal entry numbers.

Concurrent with the merge, entries are made in a new global for each record making up the pair. The entries are intended to provide a "before-merge" image. However, please note that the merge is a non-reversible process. Once the pair of records is merged, there is no automated way of undoing the process.

The application has been written to support multiple parallel jobs (threads - as specified by the site) during the merge process. However, decreased overall processing time is exchanged for increased system utilization.

Advisory Note: The merge process is a background job. Be aware that it should not be running when changes are being made to Data Dictionaries or when data conversions are taking place.

Site Parameters

The DUPLICATE RESOLUTION file (#15.1) is used to handle duplicate checking and merging duplicate record pairs for the PATIENT file (#2). It provides control over the information that is used to first identify potential duplicates within a file and then merge those verified duplicate entries.

Note: Duplicate record pairs are stored and merged in the DUPLICATE RECORD file (#15).

Several parameters are exported with this application in the DUPLICATE RESOLUTION file as default values. They are related to how potential duplicates are identified, to designating ancillary reviewers, to naming users to mail groups, and to defining the length of waiting periods applied in the process of identifying and merging duplicate records.

The following is a list of site parameter field names and descriptions. Default values that come with this application are identified. Also identified are values that the site must supply when configuring the Patient Merge application [i.e., When configuring the DUPLICATE RESOLUTION file].

Field Name	Description
MERGE MAIL GROUP	This mail group has the following three functions: <ol style="list-style-type: none"> <li data-bbox="873 1251 1386 1314">1) A bulletin is sent when a merge process has been completed, <li data-bbox="873 1325 1403 1734">2) A MailMan message is sent comprised of FROM and TO record pairs detected as having data errors during the beginning of the actual merge process. Records contained in this message are excluded from the merge process. The subject of this MailMan message is "MERGE PAIRS EXCLUDED DUE TO PROBLEMS". <li data-bbox="873 1745 1427 1892">3) A MailMan message is sent comprised of FROM and TO record pairs excluded from the merge process resulting from their

Field Name	Description
	<p>multiple relationship(s) with the record pair selected to be merged. The subject of the MailMan message is "PAIRS EXCLUDED FROM MERGE DUE TO MULTIPLE REFERENCES".</p> <p>(See the section "Schedule Process to Merge Verified Duplicates" in the <i>Duplicate Record Merge: Patient Merge User Manual</i> for more information on data error and/or multiple pair processing excluding records from being merged.)</p>
DUPLICATE MANAGER MAIL GROUP	This mail group will be sent a bulletin whenever any known problems occur during the duplicate checking or merging process. This mail group must be an entry in the MAIL GROUP file. Sites are required to supply their own mail groups.
POTENTIAL DUPLICATE THRESHOLD%	This field contains a computed percentage based on the values defined in the duplicate tests. When record pair scores evaluate equal to or above this percentage, they are considered to be potential duplicates. Record pairs may differ on the maximum score to which the percentage threshold is applied. The default potential duplicate threshold with this application release is 60%.
DAYS BEFORE FINAL VERIFY	This is the number of days within which ancillary reviewers are expected to conduct their reviews. All reviews must be resolved prior to merging the affected record pair. If action is still pending by any ancillary reviewers after this period, the site may take alternative action to:

Field Name	Description
DAYS BETWEEN VERIFY AND MERGE	<p>notify designated persons via a mail group, establish an additional reviewer, otherwise to resolve the review.</p> <p>This would only apply if verification of a potential duplicate pair were not completed before the site configurable time runs out. The number of days entered in this field should be between zero and 30. This application is exported with the default value set to five days.</p>
NUMBER OF THREADS	<p>This is the number of days to elapse between final verification of a duplicate pair and the earliest possible merge date (i.e., the date records are approved to be merged). This parameter is available for sites that may be concerned about possible on-going actions related to entries that will be merged and want to establish a period of inactivity prior to the actual merge. The value of this field is set by IRM. The number of days entered in this field should be between zero and 30. The default with this application release is five days.</p> <p>This field is used to help manage the most time consuming portion of the merge process. It is used to indicate the number of threads that should be used during the longest phase of the merge process. (Phase 3 is the longest phase of the merge process. It is described in more detail in the "Merge Process" section of this manual.) Your site can set the number of threads (jobs) according to your system resources.</p> <p>For example, an entry of 1 would indicate that only the main process would be</p>

Field Name

Description

running. An entry of 2 to 3 indicates that the main process along with one to two other threads should be used during this phase. The greater the number of threads the more parallel processing occurs. Each thread is used to process one of the more time consuming files, which reduces the amount of time the overall process takes, at the expense of increased system utilization.

The maximum allowable value is 5. It is resident in the DUPLICATE RESOLUTION file. The site parameter NUMBER OF THREADS is exported with the Patient Merge application as a default of three.

Ancillary Service Site Parameters

The necessary default values associated with each of the three ancillary services (i.e., Laboratory, Pharmacy, and Radiology) are included with this application. Sites are required to supply mail groups and/or assign individuals to receive alerts. Your site can add more ancillary services to the ancillary service multiple if desired.

The following table provides an "at a glance" view of the site parameters for all three ancillary services exported with this release. References are made to both site parameters that are default values included with this application release and to fields that require the site assignment of values. Ancillary service, field name, and default data value(s) list the information.

Ancillary Service	Field Name	Default Data Value
LABORATORY	PRIMARY FILE NUMBER	63
	SUBFILES (FIELD NUMBERS)	1;.085;2;4;5;8;
	DISPLAY NAMES FOR SUBFILES	Bld Bank;Transfusn;EM;Lab Tests;Mic
	MAIL GROUP NAME	provided by site
	INDIVIDUAL FOR ALERTS	provided by site
PHARMACY	PRIMARY FILE NUMBER	55
	SUBFILES (FIELD NUMBERS)	N/A
	DISPLAY NAMES FOR SUBFILES	N/A
	MAIL GROUP NAME	provided by site
	INDIVIDUAL FOR ALERTS	provided by site
RADIOLOGY	PRIMARY FILE NUMBER	70
	SUBFILES (FIELD NUMBERS)	2
	DISPLAY NAMES FOR SUBFILES	Radiology
	MAIL GROUP NAME	provided by site
	INDIVIDUAL FOR ALERTS	provided by site

File names and descriptions for the ancillary service site parameters are provided on the following pages. References are made to both site parameters that are default values included with this application release, and to fields that require the site assignment of values.

Field Name	Description
ANCILLARY SERVICE	Duplicate record pairs may require secondary reviews when data is present in ancillary files. This field is the name of the VISTA ancillary service associated with each duplicate record pair. Each of these ancillary services has an associated primary file, which could provide information on whether the entries are in fact, duplicates or not. In addition, the ancillary reviewers' expertise in evaluating other data in health summaries may provide information on the duplicate question. (This entry must be from three to 30 characters in length.) Three default ancillary services are released with Patient Merge: Laboratory, Pharmacy, and Radiology.
PRIMARY FILE NUMBER	Primary file number for the main file associated with the ancillary service. Enter the file number to be used for identifying duplicate record pairs for this ancillary service [e.g., LAB DATA file (#63), RADIOLOGY PATIENT file (#70), PHARMACY PATIENT file (#55), etc.]. Each ancillary service distributed with this application release contains a default value for this field.
SUBFILES (FIELD NUMBERS)	Subfiles in the primary file. To show how many particular types of data, or entries, exist for some of the subfiles, you can enter the field numbers for the subfile(s) that should be displayed. These entries should be separated by semicolons (e.g., 1;.085;2;4;5). Each ancillary service distributed with this application release contains a default value for this field.
DISPLAY NAMES FOR SUBFILES	Respective subfile names for each of the

Field Name	Description
MAIL GROUP NAME	<p>specified subfiles in the primary file. Specify the names that will be associated with the subfile numbers as they are entered in the field SUBFILES (FIELD NUMBERS). The names entered must be in the same position as their corresponding field numbers. These names will appear on the screen so you might want to make them descriptive. Semicolons (e.g., Bld Bank; Transfusn; EM; Lab Tests; Microbiol) should separate these entries. Each ancillary service distributed with this application release contains a default value for this field.</p> <p>Optional. This field is a pointer to the MAIL GROUP file. Your answer must be the name of an actual mail group entry in that file. Once potential duplicate record pairs are established from the search, populated into File #15, and reviewed by the primary reviewers, MailMan messages are automatically sent to these designated mail groups, if data is present for both entries in the potential duplicate pair in the file specified as the primary file for the ancillary service. Alerts can also be sent to individuals separately by designating one or more reviewer in the INDIVIDUAL FOR ALERTS field. Sites are required to supply their own mail groups.</p>
INDIVIDUAL FOR ALERTS	<p>Optional. Enter one or more ancillary reviewer. Once potential duplicate record pairs are established from the search, populated into File #15, and reviewed by the primary reviewers, alerts are automatically sent to these designated individuals, if data is present for both entries in the potential</p>

Field Name	Description
	duplicate pair in the file specified as the primary file for the ancillary service. Answer with the NEW PERSON NAME, INITIAL, SSN, NICK NAME, DEA#, or VA#.
	Sites are required to supply individuals for alerts. For a review by an ancillary service to be active, there must be at least one member in a specified mail group OR at least one individual specified to receive alerts.

Note: Not all ancillary services will have sub-files associated with the primary file.

New Browser Setup: XDRBROWSER1

Patient Merge brings in a new browser setup to your site. XDRBROWSER1 is specifically designed to work with Patient Merge.. It is a modified version of the VA FileMan Browser. Setup instructions for OpenM platforms are included on the following page. The following information should be entered in your TERMINAL TYPE file (#3.2) and DEVICE file (#3.5), respectively:

Entry in TERMINAL TYPE File

```
NAME: P-XDRBROWSER
  SELECTABLE AT SIGN-ON: NO
  RIGHT MARGIN: 255
  FORM FEED: #
  PAGE LENGTH: 99999
  BACK SPACE: $(8)
  OPEN EXECUTE: D OPEN^XDRDVAL
  CLOSE EXECUTE: D CLOSE^XDRDVAL
  DESCRIPTION: BROWSER FOR DUPLICATE RESOLUTION APPLICATIONS
```

Entry in DEVICE File

```
NAME: XDRBROWSER1
  $I: XDRBROW.TXT
  ASK DEVICE: NO
  ASK PARAMETERS: NO
  QUEUING: ALLOWED
  LOCATION OF TERMINAL: HFS/CRT
```

```

ASK HOST FILE: NO
ASK HFS I/O OPERATION: NO
*MARGIN WIDTH: 255
*FORM FEED: #
*PAGE LENGTH: 99999
*BACK SPACE: $C(8)
OPEN PARAMETERS: NEW:DELETE

PRE-OPEN EXECUTE: S %ZIS("HFSNAME")="XDR"_$J_".TXT" I
'$D(XDRDVALF),'$$TEST^DDBRT S %ZISQUIT=1 W $C(7),!,"Browser not selectable
from current terminal.",!

SUBTYPE: P-XDRBROWSER
TYPE: HOST FILE SERVER

```

Note: You must have a NULL device defined in the device setup.

New Browser Setup: XDRBROWSER1 (OpenM)

The following setup instructions are for OpenM platforms. The information should be entered in your TERMINAL TYPE file (#3.2) and DEVICE file (#3.5), respectively:

Entry in TERMINAL TYPE File

```

NAME: P-XDRBROWSER
SELECTABLE AT SIGN-ON: NO
RIGHT MARGIN: 255
FORM FEED: #
PAGE LENGTH: 99999
BACK SPACE: $C(8)
OPEN EXECUTE: D OPEN^XDRDVAL
CLOSE EXECUTE: D CLOSE^XDRDVAL
DESCRIPTION: BROWSER FOR DUPLICATE RESOLUTION APPLICATIONS

```

Entry in DEVICE File

```

NAME: XDRBROWSER1
$I: T:\TEMP\XDRBR.TXT (or the desired drive and directory)
ASK DEVICE: NO
ASK PARAMETERS: NO
QUEUING: ALLOWED
LOCATION OF TERMINAL: HFS/CRT
ASK HFS I/O OPERATION: NO
OPEN PARAMETERS: "NWS"

PRE-OPEN EXECUTE: S %ZIS("HFSNAME")=$P(IO, ".")_$J_"._$P(IO, ".")_2) I
'$D(XDRDVALF),'$$TEST^DDBRT S %ZISQUIT=1 W $C(7),!,"Browser not selectable from
current terminal.",!

```

Implementation and Maintenance

SUBTYPE, P-XDRBROWSER
TYPE, HOST FILE SERVER

Note: You must have a NULL device defined in the device setup.

Routine Descriptions

The following routines are distributed with Patch XT*7.3*23 (i.e., the Patient Merge application). A brief description is given for each routine.

XDR2NULL	Identifies entries with null SSN.
XDRCNT	Counts/Tallies records in file #15 by STATUS (verified, not a duplicate; potential duplicate, unverified; verified duplicate; verification in process). Also count/tally records by MERGE STATUS (not ready, ready, merged, in process). A report is generated to display the results.
XDRDADD	After two records have been matched as potential duplicates, this routine will add the two records to the DUPLICATE RECORD file.
XDRDCOMP	Compares two patients via duplicate checker algorithm.
XDRDDATA	Lists basic data on potential duplicates in global format.
XDRDEDT	Updates the STATUS field in File 15 from either Verified Duplicate or Verified Not a Duplicate to Potential Duplicate.
XDRDLIST	Prints various reports from the DUPLICATE RECORD file. It prints listings of Unverified Potential Duplicates, Ready, and Not Ready To Merge Verified Duplicates, and Merged Verified Duplicates.
XDRDMAIN	Main driver for the duplicate checking software.
XDRDPICK	Verify a pair of potential duplicates after viewing their data. Calls LOOKUP^XDRDEDT to select entries from the DUPLICATE RECORD file (#15) instead of using the standard DIC lookup.
XDRDPRE1	Generates lists of patients identified by the preliminary scan (XDRDPREL).
XDRDPREL	Preliminary identification of entries with bad data. This routine will scan the file for errors.
XDRDPRG2	Purges the XDR MERGE PROCESS file (#15.2).

Routine Descriptions

XDRDPRGE	Purges the DUPLICATE RECORD file (#15).
XDRDQUE	Starts and stops the search process. If no search is running, it allows the user to queue a search to start up. If a search has been halted, they may continue the search starting at the point they halted. Code has been added to inform user that the search can take a long time.
XDRDSHOW	Displays data in fields so users can select fields to overwrite.
XDRDSTAT	Displays the status of a search.
XDRDUP	Compares two records and makes the determination, if they are duplicates of each other.
XDRDVAL	Contains the basic code for checking whether a data value would be acceptable as input.
XDRDVAL1	Entry points for the validation of data in terms of data entry.
XDRDVAL2	(Continued from XDRDVAL1.)
XDREMSG	Error message processor.
XDRMADD	Adds entries to the DUPLICATE RECORD file.
XDRMAIN	This is currently the main driver for duplicate merge software.
XDRMAINI	Initialization routine for XDRDMAIN.
XDRMERG	Tentative update pointer nodes.
XDRMERG0	Start of non-interactive batch merge. The merge can be queued.
XDRMERG1	Tentative update pointer nodes. (Continued from XDRMERG.)
XDRMERG2	Tentative update pointer nodes. (Continue from XDRMERG.)
XDRMERGA	Start of non-interactive batch merge. (Continued from XDRMERG0.)
XDRMERGB	Tentative update pointer nodes.
XDRMVFY	Verifies potential duplicates.

XDRPREL1	Identifies problem entries which are **NOT** pointed to.
XDRPTCAN	Screens for patient records with similar names, DOBs, and SSNs.
XDRPTCLN	Compares patient's claim numbers.
XDRPTDOB	Compares patient's DOB dates.
XDRPTDOD	Compares patient's DOD dates.
XDRPTLSD	Compares last service separation dates.
XDRPTMMN	Compares mother's maiden names.
XDRPTN	Compares patient's names.
XDRPTSSN	Compares patient's SSNs.
XDRPTSX	Compares the sex of two patients.
XDRRMRG0	Triggers the data review for ancillary services - the verification process for potential duplicates that have ancillary data.
XDRRMRG1	Duplicate verification routine for RADIOLOGY.
XDRRMRG2	Retrieves patient's Health Summary.
XDRUTL	XDR routine utilities.
XDRVCHEK	Checks for entries, which have passed the number of days required for verification.

The following routines are not distributed with the Patient Merge application (e.g., Patch XT*7.3*23). They have not changed from Kernel Toolkit V. 7.3. They are, however, still necessary for the operation of Patient Merge.

XDRDADJ	Adjusts the DUPLICATE RECORD file upon merge.
XDRDFPD	Finds all potential duplicates for an entry in a file.
XDRDPDTI	When the threshold is increased, this routines will check potential duplicates against it.

Routine Descriptions

XDRDSCOR	Sets scores for duplicate checking.
XDRMPACK	Checks the PACKAGE file for special merge routines in the AFFECTS RECORD MERGE field.
XDRMRG	Merges duplicate records.
XDRMRG1	Error trap for XDRMRG.
XDRMSG	Sends various mail messages when verifying and merging have completed.
XDRU1	Contains general utilities.

The following routines have become obsolete in Kernel Toolkit with the release of the Patient Merge application:

XDRDOC	Internal documentation.
XDRDOC1	Internal documentation. (Continued from XDRDOC.)
XDRDOC2	Internal documentation. (Continued from XDRDOC.)
XDRERR	Sends a mail message, if test routine missing.
XDRHLP	Help logic for selected fields in File 15 and 15.1.
XDRPREI	Deletes DD for Files 15 and 15.1.

File List

Files and Globals

This section lists all the Patient Merge files with their file numbers, shows their global location, and gives a file description.

15 DUPLICATE RECORD FILE ^VA(15,
Data Comes with File: No

This file contains information about duplicate records in any file as defined in the two variable pointer fields: RECORD1 and RECORD2 (in the .01 and .02 fields). The status of an entry in this file may be one of the following:

Potential Duplicate, Unverified,
Verified, Not A Duplicate,
Verified Duplicate,
Verification In Process,
or Requires Resolution.

The envisioned sequence of events for this application include:

identify potential duplicates from a file,
add an entry in this file containing the two potential duplicate records,
set the STATUS field to Potential Duplicate, Unverified,
a user makes the determination of whether the two entries were truly
duplicates and changes the STATUS field appropriately.

***** W A R N I N G *****

When you add a file to a variable pointer field, FileMan deletes the input transform so save it off before the change and restore it after the change.

Cautionary Note: Please be advised that the DUPLICATE RECORD file can be used as a perpetual list of which records have been merged and when. This information could prove useful for other systems (e.g., national databases). It is recommended that this file NOT be deleted using any means other than the option provided with this package (i.e., Purge Duplicate Record File) so that data can be kept for future use. The option Purge Duplicate Record File

has a safety feature in it such that it does not allow you to purge merged duplicate records.

15.1 DUPLICATE RESOLUTION FILE ^VA(15.1,
Data Comes with File: Yes

This file is used to handle duplicate checking and merging of files that have entries in the Duplicate Record File. It is meant to provide the overall control information that would be used to first identify duplicates within a file (e.g., PATIENT file), and then to merge the entries.

15.2 XDR MERGE PROCESS FILE ^VA(15.2,
Data Comes with File: No

When a merge process is set up, all its information is stored in this file. Once a merge process has completed, that entry may be purged using the Purge Merge Process File option in the Manager Utilities menu.

15.3 XDR REPOINTED ENTRY FILE ^VA(15.3,
Data Comes with File: No

This file is used to record the entry number of the FROM record that is merged into the TO record. This can be used for FileMan to determine which entries were merged, so the IEN of the FROM record will not be reused.

15.4 MERGE IMAGES ^XDRM(
Data Comes with File: No

File #15.4 stores a pre-merge image of each individual record making up the duplicate record pair merged. This includes pointer values. This global is automatically populated immediately before the actual merge.

Templates

Print Templates

Print templates:

XDR BRIEF

XDR MERGED LIST

Sort Templates

Sort templates:

XDR MERGED LIST
XDR NOT READY TO MERGE LIST
XDR POTENTIAL DUPLICATE
XDR READY TO MERGE LIST

Form

Form:

XDR RESFILE FORM

Exported Options

Menus and Options

This section details the three menus comprising the Duplicate Resolution System [XDR MAIN MENU]. They are the Operations, Utilities, and Manager Utilities menus. The Operations and the Utilities menus should be made accessible to personnel who would be involved with the review and verification of duplicate records. These people must be authorized to hold the XDR security key to access these menus. The Manager Utilities menu is restricted to those personnel authorized to hold the XDRMGR security key. Care should be used in granting the XDRMGR security key since other options that impact system resources (e.g., Schedule Process to Merge Verified Duplicates) are controlled by it.

The following pages list these menus, associated options, and descriptions. Personnel must be authorized to hold the XDR security key to access these menus.

Operations

[XDR OPERATIONS MENU]

This menu contains options for reviewing and verifying duplicate record status, and for approving verified duplicate record pairs for merging. As was mentioned previously, it should be made accessible to personnel who would be involved with the review and verification of duplicate records.

The following is a list of the options and descriptions that comprise this menu:

Approve Verified Duplicates For Merging [XDR APPROVE FOR MERGE]

This option is used by the coordinator, or team, responsible for duplicate record verification to give final approval for selected duplicate pairs to be included in the next merge process.

Verify Potential Duplicates [XDR VERIFY ALL]

This option is used to either mark potential duplicate pairs as an actual duplicates or mark a potential duplicate pair as verified that they are not duplicates. The FROM and TO records are identified and all top level patient file fields are resolved.

Ancillary Data Review

[XDR ANCILLARY REVIEW]

Ancillary service representatives can be designated as members of mail groups notified of potential duplicates through MailMan messages. These representatives use this option to verify that record pairs are potential duplicates.

Utilities

[XDR UTILITIES MENU]

This menu contains the utility options available with the Patient Merge application. As was mentioned previously, it should be made accessible to personnel who would be involved with the review and verification of duplicate records.

The following is a list of the options and descriptions that comprise this menu:

Add Verified Duplicate Pair

[XDR ADD VERIFIED DUPS.]

This option allows the user to add a pair of records that are not already identified as potential or verified duplicate record pairs to the DUPLICATE RECORD file. The record pairs will go through comparisons (Duplicate Tests). The comparison results are computed into a value based on the similarity of one record to the other. The resulting value is measured against the Potential Duplicate Threshold Percentage. When the duplicate record pairs scores evaluate equal to or above this percentage, they are considered to be potential duplicates, and are placed in the DUPLICATE RECORD file.

Users who hold the XDRMGR key have the option to bypass the Potential Duplicate Threshold Percentage, thereby adding the duplicate record pair directly to the DUPLICATE RECORD file.

Check Merge Process Status

[XDR CHECK MERGE PROCESS STATUS]

This option indicates the status of a selected merge process (or all of them). The information is displayed by the last checkpoint during its operation. This information

includes the file that is being processed, which stage it is in, and the last internal entry processed.

Check Pair of Records to see if Duplicates [XDR CHECK PAIR]

This option allows an operator to run two records through the duplicate resolution software to evaluate their Match Score. It does not add records to the DUPLICATE RECORD file.

Display Search Status [XDR DISPLAY SEARCH STATUS]

This option displays the search status. The search status can be either Running, Halted, Error(Stop) or Completed. If you are viewing the status to check if the search is still in progress, you must check to see that the status indicated is Running, and that the Count (the current number of records checked that have been for duplicates) is steadily increasing. If the Count is not increasing, notify your Site Manager.

Edit the Status Field of a Duplicate Record [XDR EDIT DUP RECORD STATUS]

This option is used to edit the STATUS field of duplicate record entries in the DUPLICATE RECORD file. It would be used when a record pair was identified as Verified Duplicate or Verified Not A Duplicate and you want to change the STATUS back to Potential Duplicate, Unverified.

Find Potential Duplicates for an Entry in a File [XDR FIND POTENTIAL DUPLICATES]

This option will find all potential duplicates for a record entry in a selected file. The potential duplicate record pairs are then added to the DUPLICATE RECORD file. The user is prompted to enter the selected file and the record entry within that file.

Identify Potential Merge Problems [XDR VALID CHECK]

This option has been designed to display any data errors that might exist in a patient's record. Although

distributed by another package, the DG ELIGIBILITY security key is required to edit any data fields in error, retrieved and displayed by this option.

Many of the "broken database/bad data" events that would have upset an actual merge can be detected through this option. The intent is to give the site the opportunity to make any necessary edits prior to merging duplicate record pairs. Errors that are not resolved prior to the merge process will result in a pair not being merged.

Print List of File Duplicates [XDR PRINT LIST]

This option prints a selected list of duplicate record entries in a file or the files entire contents. You can choose to print Potential Duplicates, Verified Ready To Merge Duplicates, Not Ready To Merge Verified Duplicates, and Merged Verified Duplicates. You can choose to print a brief listing or a captioned listing.

Scan Possible Duplicates [XDR SCAN POSSIBLE DUPLICATES]

This option provides a rapid scan of possible duplicate record pairs by listing the zero nodes for each record entry in the selected file. If the output is not queued to a printer, the data will be sent to the FileMan Browser.

Tally STATUS and MERGE STATUS fields [XDR TALLY STATUS FIELDS]

This option produces a summary report of the number of records that are: Verified Duplicates, Verified Not Duplicates, or Potential Duplicates, Unverified. The total number of records Merged and Ready To Be Merged will be displayed.

View Duplicate Record Entries [XDR VIEW DUPLICATE RECORD]

This option is used to view duplicate record entries in a captioned format.

Manager Utilities <Locked with XDRMGR> [XDR MANAGER UTILITIES]

This menu is comprised of the various manager utilities that give IRM personnel control over critical decision elements involved with the merge process. They are designed to assist with editing site parameters, identifying duplicate records in the primary file, merging verified duplicate records, and purging files involved with the merge process that are specific to Patient Merge. Only authorized IRM personnel holding the XDRMGR security key can use this menu.

The following is a list of these options and descriptions:

Edit Site Parameters [XDR EDIT DUP RESOLUTION FILE]

Site Managers can set up patient identification and merge capability for the PATIENT file (#2) by entering the desired data into the DUPLICATE RESOLUTION file (#15.1). This option allows Site Managers to:

Set up ancillary services to review ancillary data involved with the patient identification and merge processes.

Assign reviewers to mail groups to receive the various bulletins involved with the patient identification and merge processes.

Assign reviewers to alerts designating them as recipients of pending ancillary service review activity involved with the patient identification and merge processes.

Additionally, you can use this option to edit the default site parameters in the DUPLICATE RESOLUTION file that come with this application release.

List File Entries Identified in Preliminary Scan [XDR PRELIMINARY
SCAN LIST]

This option is used to generate a list of file entries with missing zero nodes, having a bad SSN values, or missing one or more of the identifiers in the file.

Preliminary Scan of File for Errors [XDR PRELIMINARY SCAN]

This option allows you to scan through a selected file before a search to check for entries which are missing identifiers (entries presumed to be significant data elements) or have other problems.

Purge Duplicate Record File [XDR PURGE]

This option is used to purge entries in the DUPLICATE RECORD file. You can purge Potential Duplicates only, Verified Non-Duplicates, or both. The Site Manager should only use this option.

Purge Merge Process File [XDR PURGE2]

This option is used to purge all selected entries in the MERGE PROCESS file. The Site Manager should only use this option.

Restart a Merge Process [XDR RESTART MERGE PROCESS]

This option is used to schedule the restart of a merge process at the current time or at some point in the future.

Schedule Process to Merge Verified Duplicates [XDR MERGE READY
DUPLICATES]

This option merges all entries, which are ready to merge verified duplicate pairs, not included in another merge process.

Start/Halt Duplicate Search [XDR SEARCH ALL]

This utility searches a selected file for potential duplicate record pairs. It provides a choice of two methods. A Basic search starts at the beginning of the file and checks each record against a selected subgroup of potential duplicates.

Exported Options

```

Problems [XDR VALID CHECK]
-----PRNT Print List of File Duplicates
          [XDR PRINT LIST]
-----SCAN Scan Possible Duplicates [XDR
          SCAN POSSIBLE DUPLICATES]
-----TSF Tally STATUS and MERGE STATUS
          Fields [XDR TALLY STATUS
          FIELDS]
-----VIEW View Duplicate Record Entries
          [XDR VIEW DUPLICATE RECORD]

-XDRM Manager Utilities [XDR MANAGER -----EDT Edit Site Parameters [XDR EDIT
UTILITIES]                               DUP RESOLUTION FILE]
**LOCKED: XDRMGR**
-----LIST List File Entries Identified
          in Preliminary Scan [XDR
          PRELIMINARY SCAN LIST]
-----SCN Preliminary Scan of File for
          Errors [XDR PRELIMINARY SCAN]
-----PRGE Purge Duplicate Record File
          [XDR PURGE]
-----PRGM Purge Merge Process File [XDR
          PURGE2]
-----RES Restart a Merge Process [XDR
          RESTART MERGE PROCESS]
-----SCH Schedule Process to Merge
          Verified Duplicates [XDR MERGE
          READY DUPLICATES]
-----SRCH Start/Halt Duplicate Search
          [XDR SEARCH ALL]
-----STOP STOP an Active Merge Process
          [XDR STOP MERGE PROCESS]
```

Options No Longer Referenced in Duplicate Resolution Utilities

The following three options have become obsolete in Kernel Toolkit with the release of the Patient Merge application:

1. Automatically Merge all Ready Verified Duplicates [XDR AUTO MERGE]
2. Merge Selected Verified Duplicate Pair [XDR MERGE SELECTED PAIR]
3. Verify Selected Potential Duplicate Pair [XDR VERIFY SELECTED PAIR]

Security Keys

Three security keys are involved with the Patient Merge release. Their purpose is to set a layer of protection on the range of available computing capabilities. Two of which (i.e., XDR and XDRMGR) are used in the Duplicate Resolution System menu. The accessibility of this menu is based on the level of system access granted to each user. These security keys are described below:

1. The **XDR** security key authorizes access to the Duplicate Resolution System menu on the Application Utilities menu.
2. The **XDRMGR** security key is required in order to access the Manager Utilities menu. This key should be given to IRM personnel responsible for managerial control over the file being searched for duplicate record pairs and the subsequent merge process.

The **XDRMGR** security key also authorizes the holder to add a pair of records directly into the DUPLICATE RECORD file through the option Add Verified Duplicate Pair, bypassing the Potential Duplicate Threshold Percentage.

Note: Care should be used in granting the **XDRMGR** security key, since other options that impact system resources (e.g., Schedule Process to Merge Verified Duplicates) are controlled by it.

3. Although distributed by another package, the person initiating the merge process must be authorized to hold the **DG ELIGIBILITY** security key.

The **DG ELIGIBILITY** security key, although not required to select and run the option Identify Potential Merge Problems, is required to provide valid error checking results. The key may or may not be required to actually make the necessary edits.

Exported Options

Archiving and Purging

Archiving

There are no application-specific archiving procedures or recommendations for the Duplicate Record Merge: Patient Merge application.

Purging

Two options under the Manager Utilities menu are provided to facilitate the purging of Patient Merge files:

1. Purge Duplicate Record File will purge entries in the DUPLICATE RECORD file, and
2. Purge Merge Process File will purge entries in the XDR MERGE PROCESS file.

(For more detailed descriptions of these options, refer to the "Exported Options" section of this manual.)

Callable Routines

There are no callable routines in the Patient Merge application.

External Relations

Platform Requirements

The Patient Merge application requires a standard **VISTA** operating environment in order to function correctly. Check your **VISTA** environment for packages and versions installed.

Minimum application requirements are:

Kernel V. 8.0
Kernel Toolkit V. 7.3
VA FileMan V. 21.0
NDBI V. 1.0

VISTA patch installations needed to install the Duplicate Record Merge: Patient Merge application are:

Package	Designation
Immunology Case Registry	IMR*2.1*2
Inpatient Medications	PSJ*4.5*43
Integrated Billing	IB*2*94
Integrated Billing	IB*2*97
Lab Service	LR*5.2*205
Registration	DG*5.3*156
VA FileMan	DI*21*41
VA FileMan	DI*21*42

VISTA packages needed to fully utilize the capabilities of the Duplicate Record Merge: Patient Merge application are:

Patient Information Management System (PIMS) V. 5.3
Health Summary V. 2.5

The following **VISTA** packages have files that require special processing during the merge process. Additionally, a routine for National Database Integration is also included for special processing during the merge process.

Lab Service V. 5.2
Immunology Case Registry V. 2.0
Integrated Billing V. 2.0

National Database Integration (Consolidation Site activity)

Records in these files contain fields affected by the merging of the PATIENT file (#2), but are not identified as pointer fields. Routines for processing these files are sent with this application since the PATIENT file commonly points to these files. Your site can create additional routines for processing other files pointed to the PATIENT file. Any time an additional routine is created, it is necessary that an entry be made in the PACKAGE file (#9.4) in the AFFECTS RECORD MERGE subfile.

Database Integration Agreements (DBIA)

To obtain the current list of DBIAs that the Toolkit package is a custodian of:

1. Sign on to the Forum system (forum.va.gov).
2. Go to the DBA... menu.
3. Select the Integration Agreements Menu.
4. Select the Custodial Package Menu.
5. Choose the ACTIVE by Custodial Package option.
6. When this option prompts you for a package, enter TOOLKIT.
7. All current DBIAs for which TOOLKIT is custodian of are listed.

To obtain detailed information on a specific integration agreement:

1. Sign on to the Forum system (forum.va.gov).
2. Go to the DBA... menu.
3. Select the Integration Agreements Menu.
4. Select the Inquire option.
5. When prompted for "INTEGRATIONREFERENCES:", enter the integration agreement number of the DBIA you would like to display or enter the package namespace (i.e., XDR) and choose from the list.
6. The option then lists the full text of the DBIA you requested.

To obtain the current list of DBIAs that the Toolkit is a subscriber to:

1. Sign on to the Forum system (forum.va.gov).
2. Go to the DBA... menu.

3. Select the Integration Agreements Menu.
4. Select the Subscriber Package Menu.
5. Choose the Print ACTIVE by Subscribing Package option.
6. When prompted "START WITH SUBSCRIBING PACKAGE:", enter TOOLKIT (in uppercase). When prompted "GO TO SUBSCRIBING PACKAGE:", enter TOOLKIT (in uppercase).
7. All current DBIAs to which the Toolkit is a subscriber are listed.

Internal Relations

All of the Patient Merge options can be invoked independently. None require any special setup in order to run successfully.

Namespace

The Patient Merge application uses the namespace XDR.

File Numbers

Files for the application are numbered as follows: 15, 15.1, 15.2, 15.3, and 15.4. These files reside in the Global ^VA.

Mail Groups

If you don't have existing mail groups to designate as recipients of the various bulletins involved with the patient identification and merge processes, your site will have to create them using MailMan V. 7.1 utilities. If your site chooses to create mail groups, they must be added to the MAIL GROUP file (#3.8).

[For information on creating mail groups, see the VA Electronic Mail System (Mailman) Technical Manual and Systems Management Guide V. 7.1.]

You can assign these mail groups to the various bulletins involved with the patient identification and merge processes using the option Edit Site Parameters in the Manager Utilities menu. You must possess the XDRMGR security key to access this option. The following are descriptions of the mail groups associated with this application:

Use the MERGE MAIL GROUP to enter the name of the mail group that you want a:

- 1) bulletin sent to when a merge process has been completed,
- 2) bulletin sent to, comprised of FROM and TO record pairs detected as having data errors during the beginning of the actual merge process. Records contained in this message are excluded from the merge process. The subject of this MailMan message is "MERGE PAIRS EXCLUDED DUE TO PROBLEMS",

- 3) bulletin sent to, comprised of FROM and TO record pairs excluded from the merge process resulting from their multiple relationship(s) with the record pair selected to be merged. The subject of the MailMan message is "PAIRS EXCLUDED FROM MERGE DUE TO MULTIPLE REFERENCES".

(See the section "Schedule Process to Merge Verified Duplicates" in the *Duplicate Record Merge: Patient Merge User Manual* for more information on data error and/or multiple pair processing excluding records from being merged.)

Use the DUPLICATE MANAGER MAIL GROUP to enter the name of the mail group that you want a bulletin sent to when an error causes the halting of a duplicate record search, or a merge process.

Use the MAIL GROUP NAME in the Ancillary Service data screen to enter members who need to be notified when their particular ancillary service needs to process duplicate record pairs.

Note: It is up to that particular ancillary service to elect who will be notified to process potential duplicates and through what means (i.e., as a member of a mail group or as a recipient of an alert). Designated ancillary reviewers must be entered during the application setup, either as members of a specified mail group or as individuals to receive alerts.

Package-wide Variables

The Patient Merge application contains no package-wide variables.

Software Product Security

Security Features

Mail Groups and Alerts

Site Managers can assign existing mail groups, or create new ones, to designate as recipients of the various bulletins involved with the patient identification and merge processes. New mail groups can be created using MailMan V. 7.1 utilities. If your site chooses to create new mail groups, they must be added to the MAIL GROUP file.

Additionally, Site Managers can assign individuals to receive alerts. This designates them as recipients of pending ancillary service review activity. Potential duplicates are established from the search, populated into the DUPLICATE RECORD file, and reviewed by the primary reviewers. After which, if there is ancillary data for the potential duplicate pairs, alerts are automatically sent to these designated individuals .

Recipients can be designated to receive alert notifications of potential duplicate record pairs, as well as belong to mail groups designated to receive MailMan message notifications for the same duplicate entries. Should recipients process alerts first, the same duplicate record pairs will not be accessible through the option Ancillary Data Review. The opposite is also true. Should the recipients process duplicate entries using the option Ancillary Data Review first, the alert goes away.

Mail groups to receive mail messages, and individuals to receive alerts can be assigned by selecting the option Edit Site Parameters in the Manager Utilities menu. You must possess the XDRMGR key to access this option.

The following are descriptions of the mail groups associated with this application:

Use the MERGE MAIL GROUP to enter the name of the mail group that you want a:

- 1) bulletin sent to when a merge process has been completed,
- 2) bulletin sent to, comprised of FROM and TO record pairs detected as having data errors during the beginning of the actual merge process. Records contained in this message are excluded from the merge process. The subject of this MailMan message is "MERGE PAIRS EXCLUDED DUE TO PROBLEMS",
- 3) bulletin sent to, comprised of FROM and TO record pairs excluded from the merge process resulting from their multiple relationship(s) with the

record pair selected to be merged. The subject of the MailMan message is "PAIRS EXCLUDED FROM MERGE DUE TO MULTIPLE REFERENCES".

(See the section "Schedule Process to Merge Verified Duplicates" in the *Duplicate Record Merge: Patient Merge User Manual* for more information on data error and/or multiple pair processing excluding records from being merged.)

Use the DUPLICATE MANAGER MAIL GROUP to enter the name of the mail group that you want a bulletin sent to when an error causes the halting of a duplicate record search, or a merge process.

Use the MAIL GROUP NAME in the Ancillary Service data screen to enter members who need to be notified when their particular ancillary service needs to process duplicate record pairs.

Note: It is up to that particular ancillary service to elect who will be notified to process potential duplicates and through what means (i.e., as a member of a mail group or as a recipient of an alert). Designated ancillary reviewers must be entered during the application setup, either as members of a specified mail group or as individuals to receive alerts.

Remote Systems

The Patient Merge application does not transmit data to any remote system or facility database.

Archiving/Purging

Archiving

There are no application-specific archiving procedures or recommendations for the Duplicate Record Merge: Patient Merge application.

Purging

Two options under the Manager Utilities menu are provided to facilitate the purging of Patient Merge files:

1. Purge Duplicate Record File (XDR PURGE)

This option is used to purge entries in the DUPLICATE RECORD file. You can purge Potential Duplicates, Verified Non Duplicates, or both. The Site Manager should only use this option.

2. Purge Merge Process File (XDR PURGE2)

This option is used to purge selected entries in the Merge Process File. The Site Manager should only use this option.

Interfacing

There are no specialized (not VA produced) products (hardware and/or software) embedded within or required by the Patient Merge application.

Electronic Signatures

There are no electronic signatures used in the Patient Merge application.

Menus

There are no options of particular interest to Information Security Officers (ISOs) in the Patient Merge application.

Security Keys

Three security keys are involved with the Patient Merge release. Their purpose is to set a layer of protection on the range of available computing capabilities. Two of which (i.e., XDR and XDRMGR) are used in the Duplicate Resolution System menu. The accessibility of this menu is based on the level of system access granted to each user. These security keys are described below:

1. The **XDR** security key authorizes access to the Duplicate Resolution System menu on the Application Utilities menu.
2. The **XDRMGR** security key is required in order to access the Manager Utilities menu. This key should be given to IRM personnel responsible for managerial control over the file being searched for duplicate record pairs and the subsequent merge process.

The **XDRMGR** security key also authorizes the holder to add a pair of records directly into the DUPLICATE RECORD file through the option Add Verified Duplicate Pair, bypassing the Potential Duplicate Threshold Percentage.

Note: Care should be used in granting the **XDRMGR** security key, since other options that impact system resources (e.g., Schedule Process to Merge Verified Duplicates) are controlled by it.

3. Although distributed by another package, the person initiating the merge process must be authorized to hold the **DG ELIGIBILITY** security key.

The **DG ELIGIBILITY** security key, although not required to select and run the option Identify Potential Merge Problems, is required to provide valid error checking results. The key may or may not be required to actually make the necessary edits.

File Security

Number	File Name	DD	WR	DEL	RD	LAYGO	AUDIT
15	DUPLICATE RECORD	@	@	@	#	@	#
15.1	DUPLICATE RESOLUTION	@	@	@	#	@	#
15.2	XDR MERGE PROCESS	@					#
15.3	XDR REPOINTED ENTRY	@					#
15.4	MERGE IMAGES	@	@	@	@	@	@

How to Generate Online Documentation

On-line documentation about the Patient Merge application may be obtained in a number of ways:

Retrieving Online Help Using Question Marks

The use of question marks at the file and field level is described in the *VA FileMan Technical Manual*. The use of question marks within the menu system invokes help about options and menus. One question mark at the top-level menu prompt displays the items available on the menu. Two question marks will show the Common Menu available to all users, as well as any secondary menu options for the current user. Locked options are displayed, if the user holds the key. Three question marks display descriptions of the options from the OPTION file. Four question marks display a help frame, if one has been associated with this option in the OPTION file. A question mark followed by the name of an option on the current menu displays a help frame, if one has been named for that option in the OPTION file.

Print Options File

The PRINT OPTION file, in the Kernel's Menu Management Menu, displays a list of namespaced options associated with the Patient Merge application. Other namespaced entries may also be retrieved from the PRINT, INPUT, and SORT TEMPLATE files, and the SECURITY KEY, FUNCTION, BULLETIN, and HELP FRAME files.

List File Attributes

This FileMan option allows the user to generate documentation pertaining to files and file structure. Utilization of this option via the "Standard" format will yield the following data dictionary information for a specified file(s): file name and description, identifiers, cross-references, files pointed to by the file specified, files which point to the file specified, input templates, print templates, and sort templates. In addition, the following applicable data is supplied for each field in the file: field name, number, title, global location, description, help prompt, cross-reference(s), input transform, date last edited, and notes.

Using the "Global Map" format of this option generates an output which lists all cross-references for the file selected, global location of each field in the file, input templates, print templates, and sort templates.

Inquire to Option File

This Menu Manager option provides the following information about a specified option(s): option name, menu text, option description, type of option, and lock, if any. In addition, all items on the menu are listed for each menu option.

To secure information about Patient Merge options, the user must specify the name or namespace of the option(s) desired (i.e., XDR).

Appendix A

Duplicate Tests and Scores: Technical Description

The following attributes from the PATIENT file (#2) are used in the identification of potential duplicate pairs. Shown below are the field names, field number, and the positive and negative weights currently assigned to these values in the determination of a numeric value for comparison with the potential total value.

Field Name	Field Number	Positive Weight (score)	Negative Weight (score)
Name	.01	100	-60
SSN	.09	100	-60
Claim Number	.313	80	-60
Date of Birth	.03	60	-40
Date of Death	.351	50	-50
Mother's Maiden Name	.2403	50	-90
Last Separation Date	.327	50	-40
Sex	.02	20	-90

Initial Screen for Potential Duplicates

The current file entry is tested against other entries in the file to determine records that might be potential duplicates. The initial screen for potential duplicates is a very broad test. It is based on Name, SSN, and Date of Birth data. Records that are identified are compared in further detail to determine whether they should be retained as potential duplicates. Other pertinent information applying to the initial screen:

If the current file entry has an SSN value beginning with five zero digits, it is skipped as a test patient entry.

Any entry in the file which has a name beginning with the same last name and first initial (as determined from the "B" cross-reference) is included.

Any entry with the same last 4 digits of the social security number (as determined from the "BS" cross-reference) is included.

Any entry with the same date of birth or with a date of birth in which the day of the month is transposed (as determined from the "ADOB" cross-reference) is included.

Further Testing of Potential Duplicates

After the initial screen, the current file entry is tested against each potential duplicate found with respect to each of the attributes listed on the previous page. If either one of the potential duplicate pair entries does not contain data for an attribute, that attribute is ignored and is not included within the final score.

Each test is associated with a positive score and a negative score. The positive score (or a fraction of it) is given, if the two entries match (or are considered a partial match). The negative score is assigned, if the two entries do not match. The scores assigned (positive or negative) for each attribute are totaled and compared to the total positive score possible for the attributes included. If the actual score is equal to or greater than a set percentage of the total possible, the pair of entries are added as a potential duplicate pair. Currently, the threshold percentage is set at 60.

Tests for Name

Each possible name for the entry being matched is used during the comparisons. The name that gives the highest score is used as the score for the pair. (Possible names include the primary name entry, any aliases assigned, and names which are generated if a ZZ string precedes a name.)

Names which match exactly are given the full positive score.

If first names match and last names match the pair is given 80% of the full score.

If first names match by soundex and last names match by soundex the pair is given 60% of the full score.

If last names and first initials match the pair is given 50% of the full score.

If last names match the pair is given 40% of the full score.

If first names match by soundex, the pair is given 20% of the full score.

All other pairs are given the negative score.

Tests for Social Security Number

If either entry in the pair has a pseudo-Social Security Number, the testing of SSNs is omitted.

Any pair in which the SSNs differ only by a single digit, or only by the transposition of two adjacent digits, are given the full positive score.

If the first initial of the last name and the last four digits of the SSN match, 80% of the full score is given.

If the last four digits of the SSN match, 60% of the full score is given.
 If the first five digits of the SSN match, 40% of the full score is given.
 If the first three digits match and at least two of the last four digits match in value and position, 20% of the full score is given.
 If the middle two digits (digits 4 and 5) match between entries and at least two of the last four digits match in value and position, 20% of the full score is given.
 Pairs which do not meet any of the above criteria are given the full negative score.

Tests for Claim Number

If the Veteran's Claim Numbers are identical, the full positive value is given.
 If the Veteran's Claim Numbers differ by only one digit or by transposition of two adjacent digits, 80% of the value is given.
 Any other pairs with claim numbers are assigned the full negative value.

Tests for Dates

The Date of Birth, Date of Death, and Last Separation Date are all processed in the same manner.

If the pair of dates match exactly, the full positive value is given.
 If the pair of dates differ only in a single digit, or by transposition of an adjacent pair of digits, 80% of the value is given.
 If one or both of the dates is missing the day of the month, but they are otherwise identical, 80% of the value is given.
 If one or both of the dates is missing the month, but they are from the same year, 60% of the value is given.
 Other date pairs are given a negative value.

Tests for Mother's Maiden Name

The data in Mother's Maiden Name is processed in the same manner as the Name field, except initial processing is performed to attempt to isolate the actual maiden name. Since a number of entries including information such as '(Deceased)' or '(Living)' have been observed, only text prior to a parenthesis is taken. If there are commas included, then only the text preceding the comma is taken. If there are no commas, but a space in the name, the name after the space is taken [this might affect some multi-string names, but the last string should be the same, if they match].

Tests for Gender

The gender values (Sex field) are tested for being the same or different. The positive value is given, if they are the same, but this is a much smaller positive value than most scores. The negative value, however, for the sex of the two entries being different is much larger, since it basically separates the two entries from each other.



**DUPLICATE RECORD MERGE:
PATIENT MERGE
TECHNICAL MANUAL**

PATCH XT*7.3*23

April 1998

Department of Veterans Affairs
VISTA Software Development
Infrastructure Product Line

Preface

This is the **Technical Manual** for the Veterans Health Information Systems and Technology Architecture (**VISTA**) [formerly the Decentralized Hospital Computer Program (DHCP)] Duplicate Record Merge: Patient Merge application. It is designed to provide you, the Site Manager/Information Resource Management (IRM) Service Chief, the necessary information for use in the technical operation of the software. It is intended for use by technical computer personnel and not designed for use by the typical end user.

The Patient Merge component of the Duplicate Record Merge software will enhance the ability to associate appropriate data with a single patient identifier. It provides the tools necessary to automatically identify patient records that have a high likelihood of being duplicates. Patient Merge is being released as a patch to Kernel Toolkit 7.3.

Preface

Contents

Introduction	1
Product Description	1
Orientation	5
Implementation and Maintenance	7
Site Parameters.....	9
Ancillary Service Site Parameters.....	12
New Browser Setup: XDRBROWSER1	16
New Browser Setup: XDRBROWSER1 (OpenM).....	17
Routine Descriptions	19
File List	23
Files and Globals.....	23
Templates	24
Form.....	25
Exported Options.....	27
Menus and Options	27
Menu and Option Diagrams	33
Options No Longer Referenced in Duplicate Resolution Utilities.....	35
Security Keys.....	35
Archiving and Purging.....	37
Archiving	37
Purging	37
Callable Routines.....	39
External Relations.....	41
Platform Requirements.....	41
Database Integration Agreements (DBIA).....	42
Internal Relations.....	45
Namespace.....	45
File Numbers.....	45
Mail Groups	45
Package-wide Variables	47
Software Product Security.....	49
Security Features	49

How to Generate Online Documentation.....	53
Retrieving Online Help Using Question Marks.....	53
Print Options File	53
List File Attributes	53
Inquire to Option File	54
Appendix A.....	55
Duplicate Tests and Scores: Technical Description.....	55
Initial Screen for Potential Duplicates.....	55
Further Testing of Potential Duplicates	56
Tests for Name	56
Tests for Social Security Number.....	56
Tests for Claim Number	57
Tests for Dates.....	57
Tests for Mother's Maiden Name	57
Tests for Gender.....	58