**25 February 2019**

CDR Immunization Table

for the MHS Data Repository (MDR)

(Version 1.02.00)

Current Specification

Revision History

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Version | Date | Originator | Para/Tbl/Fig | Description of Change |
| 1.00.00 | 8/15/2012 | J. MacLeod |  | Baseline |
| 1.01.00 | 3/26/2018 | J. MacLeod | Table 5 | Adjusted for NDAA-related changes in the LVM.  Dropped ACV.  In the LVM merge logic, corrected “Date Given” to “Date of Processing” |
| 1.02.00 | 2/25/2019 | J. MacLeod | Table 5 | Adjusted the logic for the ACV Group to have it be set to blank for dates on or after 1/1/2019. |

# MDR CDR Immunization Table

1. Background

This specification describes the process required to create the Military Health System (MHS) Data Repository (MDR) Clinical Data Repository (CDR) Immunization table based on data received from the CDR.

1. Sources

The source data files used to create the MDR Immunization tables are extracted from the Armed Forces Health Longitudinal Technology Application (AHLTA) CDR. The transfer of the raw source extracts is handled by the Defense Health Services System (DHSS) for loading into the MDR for further processing according to routine MDR operations. The source file is:

**Table 1: Source**

| **Source** | **Data Files** | **Purpose** |
| --- | --- | --- |
| CDR Immunization Data | IMIMM\*.DAT | Collection of records describing immunizations, in raw text form, prepared in accordance with the ICD. |

1. Transmission (Format and Frequency)

Source files are provided according to the frequency described in the table below.

**Table 2: Frequency of Source File**

|  |  |
| --- | --- |
| **Source File** | **Frequency** |
| CDR Immunization Table | Weekly |

1. Organization and batching

Source Data: The first step in MDR processing is to batch records received from CDR. Raw data batches are stored in /mdr/raw/cdr according to routine MDR operating procedures.

Output Products: The MDR Immunization processor outputs a single SAS data set containing all years of immunizations. The processor performs merges and field derivations, and must also apply updates to immunization records across raw data extracts. Table 3 contains the location and names of the output products. The preparation of them is described in subsequent sections of this document.

**Table 3: CDR Immunization Processor Output Product**

|  |  |  |
| --- | --- | --- |
| **MDR Immunization Processor** | **File Naming Convention** | **Member Name** |
| MDR CDR Immunization SAS data set | /mdr/pub/cdr/imm/ | immunization.sas7bdat |

Archival of files is also required, so that corresponding “apub” and other processing files (i.e. log, aprod, etc) are also loaded into the MDR according to routine operating procedures.

1. Receiving Filters

The feeds of immunization data frequently contain more than one record for a single combination of CDR Patient Id, Immunization Id, and Date Given, the three-part key for the immunization data. In some cases the records are complete duplicates with the exception of the row number. The raw data should be de-duped by the 3-part key, keeping only the most recent record where the most recent record is identified as the record with the most recent feed date, highest file number and highest row number.

Drop any records on which one or more of the key fields is missing/blank.

1. Update processes

The feeds of raw immunization data contain insert, update, and delete records. These records shall be used to maintain the master immunizations data base. First apply delete records, then insert records, and finally update records. Once the database has been updated, assign the internally-derived variables.

1. Field Transformations and File Types

This section of this functional specification describes the data merges that are necessary to append some of the fields in the MDR Immunization data. There are several merges required to prepare the MDR CDR Immunization File as described in Table 4.

**Table 4: External File Merges**

| **Merge** | **Date Matching** | **Additional Matching** |
| --- | --- | --- |
| CDR Patient Table | None | See CDR Patient Specification |
| Longitudinal VM | Date of processing | EDIPN |
| Immunization List Lookup | None | Immunization Id |

Business rules for each of the appended fields that result from the file merges and formats are described in the body of Table 5.

1. record layout and content

The CDR Immunization data are stored as one SAS dataset. The dataset file name is immunization.sas7bdat. The dataset is prepared according to the derivation rules listed in Table 5.

**Table 5: CDR Immunization SAS Data Set**

| **Variable Name** | **SAS Name** | **Format** | **Business Rule** |
| --- | --- | --- | --- |
| Immunization Id | IMMUNIZATION\_ID | 8 | No transformation. |
| Series | SERIES | 8 | No transformation. |
| Date Given | DATE\_GIVEN | 8 | No transformation. |
| Lot Number | LOT\_NUMBER | $16 | No transformation. |
| Dose | DOSE | $10 | No transformation. |
| Route | ROUTE\_TEXT | $25 | No transformation. |
| CDR Provider Id | CDR\_PROVIDER\_ID | $20 | No transformation. |
| Reaction Size | REACTION\_SIZE | $3 | No transformation. |
| Date Next Due | DATE\_NEXT\_DUE | 8 | No transformation. |
| Date Exemption Expiration | DATE\_EXEMPTION\_EXPIRATION | 8 | No transformation. |
| Reaction Result | REACTION\_RESULT | $1 | No transformation. |
| Manufacturer Code | MANUFACTURER\_CODE | $3 | No transformation. |
| Immunization Exempt Id | IMMUNIZATION\_EXEMPT\_ID | 8 | No transformation. |
| CDR Patient Id | CDR\_PATIENT\_ID | $20 | No transformation. |
| MTF Code | MTF\_CODE | $10 | No transformation. |
| Site | SITE | $11 | No transformation. |
| Feed Date | FEEDDT | $8 | Date of the feed in which this entry most recently came. YYYMMDD format. |
| **CDR Patient Table Merge** | | | |
| EDIPN | EDIPN | $10 | Obtained from merge to CDR Patient table, based on CDR\_PATIENT\_ID. |
| Sponsor SSN | SPONSSN | $9 | Obtained from merge to CDR Patient table, based on CDR\_PATIENT\_ID |
| Universal Patient Identifier | UPID | $14 | Obtained from merge to CDR Patient table, based on CDR\_PATIENT\_ID |
| Patient SSN | PATSSN | $9 | Obtained from merge to CDR Patient table, based on CDR\_PATIENT\_ID |
| Patient Date of Birth | PATDOB | 8 | Obtained from merge to CDR Patient table, based on CDR\_PATIENT\_ID |
| **LVM Merge** | | | |
| DEERS Beneficiary Category | BENCAT | $3 | Fill with DEERS beneficiary category from LVM based on EDIPN, if the date of processing is between the begin date and the end date associated with the DEERS beneficiary category. If no match for the person or the bencat is Z, set to “UNK”.  See VM6 Specification, Exhibits G-18 and 19 for segment and field positions. |
| DEERS Common Beneficiary Category | COMBEN | $1 | Derived from DEERS Beneficiary Category during LVM merge using MDR utility programs. See VM6 Specification, section A.1.12 for derivation.  If no match for the person, set to “3”. |
| DEERS ZIP Code | DEERSZIP | $5 | Fill with DEERS ZIP code from LVM based on EDIPN, if the date of processing is between the begin date and the end date associated with the DEERS ZIP code. If no match for the person, set to blank.  See VM6 Specification, Exhibits G-18 and 19 for segment and field positions. |
| DEERS Enrollment DMIS Id | DENRSITE | $4 | Fill with enrollment DMISID from LVM based on EDIPN, if the date of processing is between the begin date and the end date associated with the enrollment site. If no match for the person, set to blank.  See VM6 Specification, Exhibits G-18 and 19 for segment and field positions. |
| DEERS Sponsor Service | DSPONSVC | $1 | Fill with DEERS sponsor service from LVM based on EDIPN, if the date of processing is between the begin date and the end date associated with the DEERS sponsor service. If no match for the person, set to blank.  See VM6 Specification, Exhibits G-18 and 19 for segment and field positions. |
| DEERS Sponsor Service Aggregate | DSVCAGG | $1 | Fill with DEERS sponsor service (aggregate) from LVM based on EDIPN, if the date of processing is between the begin date and the end date associated with the DEERS sponsor service (aggregate). If no match for the person, set to blank.  See VM6 Specification, Exhibits G-18 and 19 for segment and field positions. |
| DEERS Gender | GENDER | $1 | Fill with gender from LVM based on EDIPN. If the gender is blank or U, set to “Z”. |
| DEERS Enrollment HCDP | HCDP | $3 | Fill with DEERS HCDP code from LVM based on EDIPN, if the date of processing is between the begin date and the end date associated with the DEERS HCDP code. If no match for the person, set to blank.  See VM6 Specification, Exhibits G-18 and 19 for segment and field positions. |
| DEERS Eligibility Group | ELG\_GRP | $2 | Fill with Eligibility Group from LVM if the date of processing is between the begin date and the end date of the associated segment. If no match for the person, set to “Z”.  See VM6 Beneficiary Specification, Exhibit G19. |
| DEERS Enrollment Group | ENR\_GRP | $2 | Fill with Enrollment Group from LVM if the date of processing is between the begin date and the end date of the associated segment. If no match for the person, set to “Z”.  See VM6 Beneficiary Specification, Exhibit G19. |
| DEERS PCM Type | PCM\_TYPE | $1 | Fill with PCM Type from LVM if the date of processing is between the begin date and the end date of the associated segment. If no match for the person, set to “Z”.  See VM6 Beneficiary Specification, Exhibit G19. |
| DEERS Assigned Health Care Delivery Program Code | HCDP\_ASSGN | $3 | Fill with Assigned HCDP Code from LVM if the date of processing is between the begin date and the end date of the associated segment.  See VM6 Beneficiary Specification, Exhibit G19. |
| DEERS ACV Group | ACVGROUP | $2 | For dates on or after 1/19/2019:  Set to blank.  For dates prior to 1/19/2019:  Derived by the MDR utilities during LVM merge based on Enrollment Group, PCM Type, Eligibility Group, and Common Beneficiary Category or ACV and Common Beneficiary Category depending on whether date of processing is before or after 1/1/18. If no match for the person, set to “O”.  See VM6 Beneficiary Specification, Section G.3 for details. |
| **Immunization List Lookup Table Merge** | | | |
| Immunization | IMMUNIZATION | $50 | Merge to the Immunization List Lookup table by Immunization Id. |
| **Internally-Derived Fields** | | | |
| Fiscal Month | FM | $2 | Fiscal month of Date Given |
| Fiscal Year | FY | $4 | Fiscal Year of Date Given |
| Patient Age | PATAGE | 3 | Derive from Patient Date of Birth and Date Given. |
| Age Group Code | AGEGRP | $1 | If Patient Age is missing or less than 0, set to Z. Otherwise set according to the following table:   | **Patient Age** | **Age Group** | | --- | --- | | 0 to 4 | A | | 5 to 14 | B | | 15 to 17 | C | | 18 to 24 | D | | 25 to 34 | E | | 35 to 44 | F | | 45 to 64 | G | | 65+ | H | |

1. Refresh Frequency

Frequency of updates:

* Weekly.

1. Data Quality

It is expected that when the CDR Immunization processor is run each week, that basic quality checks are performed throughout the process. It is recommended that the DHSS vendor develop a spreadsheet which tracks key characteristics of the data across processing cycles; making it relatively easy to understand how the data should generally look. DHSS vendors need to review these statistics each month prior to releasing the data. Defense Health Cost Assessment and Program Evaluation (DHCAPE) (the functional proponent and the specification author) should be contacted immediately should any quality issues arise. These checks, at a minimum, should include:

* Total record counts in the data feed should have a relatively stable distribution across FY and FM. Any anomalies should immediately be investigated.
* The percentage of records ‘cleaned out’ each processing cycle should be similar in scope and proportion across processing cycles.
* The number of records that match when doing the CDR Patient table merge should be consistent.
* The distribution of all categorical fields (ex. Reaction Result, Immunization) should be consistent. The results of proc freq analyses will verify this.
* The number of null values for important fields such as CDR\_PATIENT\_ID, Date Given, and Immunization Id should be tracked across monthly updates.
* When reading in the immunization data feed, a small number of records should be printed off and manually inspected to ensure they have read in properly and the percentage of records that are deletes, inserts, and updates should be compared for consistency across processing cycles.
* Cross tabulations should be reviewed on derived elements to ensure the derivation logic works.
* A data flow tracker should be built to ensure that all records that are intended to make it into the final CDR Immunization dataset do. In other words, all inserts, updates, and deletions should be tracked and explained in the data flow worksheet.