"README.DOC" FILE FOR THE
QUARTERLY DATA EXTRACT (QDE) FROM THE
FDA ADVERSE EVENT REPORTING SYSTEM (FAERS)

U.S. FOOD AND DRUG ADMINISTRATION (FDA)
CENTER FOR DRUG EVALUATION AND RESEARCH (CDER)
OFFICE OF SURVEILLANCE AND EPIDEMIOLOGY (OSE)

REVISED June 2015

IMPORTANT:
This document describes significant changes resulting from FDA's transition from Legacy AERS (LAERS) to the new FDA AERS (FAERS) database at FDA. We have added data to the FAERS database structure and have made minor changes to existing field contents. Users of the Quarterly Data Extract (QDE) should review all of the changes to the new extract before loading the files into their systems.

TABLE OF CONTENTS

A. INTRODUCTION

You are reading the README.DOC file that accompanies the Quarterly Data Extract from the FDA Adverse Event Reporting System (FAERS). FAERS is a computerized database for the spontaneous reporting of adverse events and medication errors involving human drugs and therapeutic biological products. FAERS began on September 10, 2012, and replaced the Adverse Event Reporting System (also known as Legacy AERS), which was decommissioned on August 27, 2012. (For more information on FAERS, please see:

http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm)

Each extract covers reports received by FAERS during one quarter of the Year. An exception is the first FAERS extract which was a non-standard quarter from August 28, 2012 to December 31, 2012, the time period from the decommissioning of LAERS until the end of the fourth quarter of 2012. The first extract also included Periodic (Non-Expedited) reports from the July/August 2012 time period.

The data are provided in two distinct formats in the extract:

1. ASCII files, in which data elements are separated from each other by a '$' sign ("$ delimited"). Please refer to the ASC_NTS.DOC file for additional information on this format.

2. XML files conforming to the guidelines of the International Conference on Harmonisation (ICH) concerning transmission of Individual Case Safety Reports (ICSR). Please refer to the XML_NTS.DOC file for additional information on this format.

Although an effort has been made to make the ASCII consistent with the XML output file, some of the data elements represented in XML are not represented in ASCII, and a few of the data elements represented in ASCII are not represented in XML. This is due to the very different natures of the two formats and to the fact that ICH E2b/M2 specifically defines the allowable data elements for XML. Furthermore, neither the ASCII nor the XML is intended to include all possible data fields.

If you wish to obtain individual case reports through the Freedom of Information Act (FOIA), please refer to these reports by the case report number only. The request must identify each case report you are interested in receiving. The case report number is included in the field/data element labeled: (1) “CASE” in the AERS ASCII format, (2) "CASEID" in the FAERS ASCII format (data files starting the 4th quarter 2012), or (3) “safetyreportid” in the SGML/XML format of the quarterly data files.

B. CLINICAL CAVEATS

Disclaimer: Submission of a safety report does not constitute an admission that medical personnel, user facility, importer, distributor, manufacturer or product caused or contributed to the event. The information in these reports has not been scientifically or otherwise verified as to a cause and effect relationship and cannot be used to estimate the incidence of these events.

There are some important things to keep in mind when reviewing or analyzing FAERS data:

- For any given report, there is no certainty that a suspected drug caused the reaction. This is because physicians are encouraged to report suspected reactions; however, the event may have been related to the underlying disease being treated, or caused by some other drug being taken concurrently, or simply occurred by chance at that time.

- Accumulated reports cannot be used to calculate incidence (occurrence rates) or to estimate drug risk. Comparisons between drugs cannot be made from these data.

- The Event code provided in the Reaction file DRUG_REC_ACT column in the ASCII extract is provided to indicate that the specific Event reoccurred when one of the Suspect Drugs was reintroduced. The listing of this Event can therefore be associated with any of the listed Suspect drugs in the case.

  NOTE: Cases submitted in Paper format (e.g., MedWatch 3500A) will not display an Event code for this DRUG_REC_ACT ASCII data element; instead QDE users should use the Rechallenge value “Y” in the DRUG file as an indicator that the reactions listed in the case reoccurred when one of the Suspect drugs was reintroduced. Cases submitted in E2B format

will provide specific Events that reoccurred (Positive Rechallenge) when one of the Suspect drugs was reintroduced.

- The Event code provided in the <drugrecuraction> tag is displayed in the XML extract along with the Drug group according to the ICSR E2B DTD specifications; however, the placement of this Event code along with the Suspect Drug does not imply association with the specific drug as the event can be associated with any of the Suspect drugs listed in the case.

  NOTE: Only the tag for <drugrecurreadministration> will display a "1" when a Positive Rechallenge has reoccurred inferring that one or more of the reactions listed in the case has reoccurred when one of the Suspect Drugs was reintroduced.

C. CAVEATS for Users who are Converting from LAERS to FAERS

The following notes are intended to assist users in transitioning their IT systems from LAERS to FAERS. Users should take into account the following IT and data changes while importing the new FAERS QDE files:

While the previous LAERS database was Individual Safety Report (ISR) based, the new FAERS database is Case/Version based. In LAERS, a Case consisted of one or more ISRs (that is, Initial and Follow-up reports), and each ISR number represented a separate version of a case. So a case could contain multiple ISRs within it, and the latest ISR represented the most current information about a particular case (for example, Follow-up 1 would have the most up-to-date information about a case containing an Initial ISR and a Follow-up 1 ISR). However in FAERS, the ISR concept is no longer being employed, and instead, each unique submission of a case received is given a version number (for example, Case# 1234567, Version 1). The first Version received (formerly called Initial) will be version 1; the second Version received (formerly called Follow-up 1) will be version 2, and so on. The latest version of a case represents the most current information about a particular Case. As a result, the FAERS QDE output files will always provide the latest, most current, Version of a Case available at the time the QDE is run.

It is possible that a case being updated during a specific quarter to add a subsequent version to the Case will not appear in the QDE output for that particular quarter if it is not yet completed when the QDE output files are prepared. This case will appear in a subsequent output, most likely the following quarter. We do not expect this scenario to affect a significant number of cases each quarter. Another scenario may occur where a version 2 of a case may result in not receiving a version 1 of the case. This would occur if version 1 and version 2 are both received and updated within a quarter. In that case, only the latest, most current, version 2 will appear in the QDE extract. We expect this scenario to happen infrequently.

In exceptional circumstances, cases may not show up in the extract in which they are expected. The reason is the switch to a Case/Version based system and the version update process. In this situation the cases will show up in the next QDE extract. In each version of a case, users can search for when the initial version of a case was received by FDA by using the <receiptdate> XML tag or the INIT_FDA_DATE ASCII tag.

While Legacy extracts provided prior to FAERS forced day or month data (for example, 201202 = 20120201 or 2012 = 20120101) where none was provided, FAERS extracts will provide only the exact data submitted in the case (for example, 201202 = 201202 or 2012 = 2012). Therefore, partial dates may be present in certain date fields (for example, in XML: drugstartdate, drugenddate – in ASCII: START_DT, END_DT). We did this to increase data quality, as previously there was a loss of accuracy in some scenarios.

In the old Legacy extract, we had a field DOSE_VBM containing the verbatim text for dose, frequency, and route (if provided), exactly as they were entered on the report. In the new FAERS extract, we have added new fields (that is, DOSE_AMT, DOSE_UNIT, DOSE_FREQ, and ROUTE) that capture the individual data elements for the dose, frequency, and route and will be provided in addition to the DOSE_VBM field. The DOSE_VBM field will only be displayed if data cannot be completely captured in the separate distinct dose, frequency, and route fields. It should be noted that in the sample data provided for testing, the old LAERS data was run using the new FAERS logic and will display data in the DOSE_VBM field as this was how the data was originally collected. However in the future extracts, the DOSE_VBM field will not be used as frequently.

In general we made the new extract the same as the old; however these are some of the few exceptions. It is important to note that the CONFID field will no longer be included in the ASCII extract because no matter what is provided in the Case, FDA does not share this information with the public. We also removed the IMAGE field. We removed the field DEATH_DT (Date patient died) to comply with privacy guidelines. You should account for these changes if you were previously importing these data elements into your database.

The transition to the new FAERS dictionaries will result in some minor data changes in the Medicinal Product values. For example, in the <medicinalproduct> tag the FAERS dictionary upgrade may result in minor differences in values in some cases (for example, LAERS=BUPROPION HCL, FAERS= BUPROPION HYDROCHLORIDE).

In the ASCII AGE field, users may occasionally encounter large numeric age values because some manufacturers submit the patient age in DAYS, even for adult patients. For example, a 30-year (YR) old will be displayed as a 10,950-day (DY) old. Therefore, users should note the AGE_UNIT accompanying the large AGE value and account for ages not listed with the year (YR) age unit.

In the XML file, the <duplicatesource> tag will display the coded name of the manufacturer sending report, if available. If there is no coded name, we will display the verbatim name of organization sending the case. Similarly, in the ASCII MFR_SNDR field, the coded name of the manufacturer sending the case will be displayed, if available. If there is no coded name, we will display the verbatim name of the organization sending the case. We are doing this to promote consistency between the ASCII and XML file outputs.

In the XML and ASCII extracts, the new XML tag <occurcountry> and the new ASCII field OCCR_COUNTRY, respectively, capture the country where the reaction/event occurred. The new tag and field will allow you to distinguish between domestic and foreign cases.

Additionally, Report Source data counts will be lower in FAERS as compared with LAERS counts because FAERS only gives the latest data of a case. In the past, if versions contained multiple report sources all would be captured. Under FAERS, only the latest version report sources are captured.

D. HOW THE FILES ARE ORGANIZED

The Quarterly Data Extract contains the following:

1. ASCII, which contains the ASCII data and informational files.

2. XML, which contains the XML data and informational files.

3. README.DOC, the informational file you are now reading.

4. FAQs – Frequently Asked Questions.

E. FILE NAME NOMENCLATURE

In the ASCII format, file names have the format <file-descriptor>yyQq, where <file-descriptor> is a 4-letter abbreviation for the data source, 'yy' is a 2-digit identifier for the year, 'Q' is the letter Q, and 'q' is a 1-digit identifier for the quarter. As an example, the ASCII demographic file for the 4th quarter of 2012 is represented as DEMO12Q4.

(The set of seven ASCII data files in each extract contains data for the full quarter covered by the extract. These files contain demographic, drug, reaction, outcome, report source, drug therapy dates, and Indication information, respectively.)

In the XML format, file names have the format ADRyyQq, where 'yy' is a 2-digit identifier for the year, 'Q' is the letter Q, and 'q' is an identifier for the quarter. As an example, the XML file for fourth quarter of 2012 is named ADR12Q4.

F. ASCII FILES

ASCII Data Files:

The ASCII data files are '$' delimited; that is, a '$' is used to separate the data fields. These files can be imported into SAS or into Access or other database program. Some data files (especially DRUGyyQq and REACyyQq) may exceed the maximum number of records that can be imported into spreadsheet programs such as earlier versions of MS Excel (for example in Office 2003, Excel has a limit of 64,000 rows).

1. DEMOyyQq.TXT contains patient demographic and administrative information, a single record for each event report.

2. DRUGyyQq.TXT contains drug/biologic information for as many medications as were reported for the event (1 or more per event).

3. REACyyQq.TXT contains all "Medical Dictionary for Regulatory Activities" (MedDRA) terms coded for the event (1 or more). For more

information on MedDRA, please contact: TRW, VAR 1/6A/MSSO, 12011 Sunset Hills Road, Reston, VA 20190-3285, USA; website is www.meddramsso.com

4. OUTCyyQq.TXT contains patient outcomes for the event (0 or more).

5. RPSRyyQq.TXT contains report sources for event (0 or more).

6. THERyyQq.TXT contains drug therapy start dates and end dates for the reported drugs (0 or more per drug per event).

7. INDIyyQq.TXT contains all "Medical Dictionary for Regulatory Activities" (MedDRA) terms coded for the indications for use (diagnoses) for the reported drugs (0 or more per drug per event).

NOTE: The MedDRA version used in the Reaction and Indication files is the current one from MSSO (www.meddra.org) and upversioned per CDER guidelines. MedDRA codes are updated biannually and version information can be determined by reviewing the <reactionmeddraversionpt> tag in the XML.

ASCII Informational Files:
---

1. ASC_NTS.DOC shows in some detail the organization and content of the ASCII data files.

2. XXXXyyQq.PDF gives null (that is, no data) counts and frequency counts for selected fields in the ASCII datasets. (The frequency counts also include the number of null values.)

G. XML FILES

XML Data File:

ADRyyQq.XML, data in a format generally conforming to the standards of the International Conference on Harmonization (ICH) for the transmission of adverse reaction reports (E2b) and its technical implementation (M2). While the previous LAERS XML files were provided in 3 monthly increments, the new FAERS XML extract will contain the entire quarter of available data in one file.

XML Informational File:
---

1. XML_NTS.DOC, which gives an introduction to the E2b standard and detailed notes about its data field descriptions.

2. XMLyyQq.PDF provides a total Case count.

H. REVISION HISTORY

April 2013 (fourth quarter 2012 data)

The first extract from FAERS contained data from August 28, 2012 to December 31, 2012. LAERS was shut down on August 27, 2012, and FAERS was deployed on September 10, 2012. Because of this transition, the initial extract is slightly larger.

For LAERS revision history details, refer to files from previous extracts, available at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/ucm082193.htm.

Jan - Mar (Q1), 2013

No Changes

Apr - Jun (Q2), 2013

No Changes

Jul - Sep (Q3), 2013

No Changes

Oct - Dec (Q4), 2013

No Changes

Jan - Mar (Q1), 2014

No Changes

Apr - Jun (Q2), 2014

No Changes

Jul - Sep (Q3), 2014

Minor corrections were made to sections C & F to improve readability and ensure current links.

Additionally, new tags were added to the 2014Q3 XML and ASCII extract files to provide QDE users with a more robust dataset to perform analysis.

The new tags implemented with this (2014Q3) release are as follows:

- Added new field for Authority Number (AUTH_NUM) in ASCII Demographic file and new tag (<authoritynumb>) added to XML extract populated with Regulatory Authority's case report number, when available.
- Added new field for Literature Reference (LIT_REF) in ASCII Demographic file and new tag (<literaturereference>) added to XML extract populated with Literature Reference information, when available.
- Added new field for Age Group (AGE_GRP) field in ASCII Demographic file and new tag (<patientagegroup>) added to XML extract populated with Age Group code, when available, as follows:

CODE	MEANING_TEXT
N	Neonate
I	Infant
C	Child
T	Adolescent
A	Adult
E	Elderly

- Added new field for Product Active Ingredient (PROD_AI) in ASCII Drug file and new tag (<activesubstancename>) added to XML extract populated with Product Active Ingredient, when available.
- Added new field for Drug Recur Action (DRUG_REC_ACT) in ASCII Reaction file and new tag (<drugrecuraction>) added to XML extract populated with the Reaction/Event information if/when Rechallenge equals Y (Positive Rechallenge).
- Added new tag (<narrativeincludeclinical>) in XML extract populated with Case Event Date, when available. NOTE: This tag does NOT include Case Narrative.

Modified field header from GNDR_COD to SEX in ASCII Demographic file.

See specific detailed information about each new data element in the XML_NTS.doc and ASC_NTS.doc files.

Oct - Dec (Q4), 2014

Modified Section B – Clinical Caveats to provide important details that should be noted when using the data in the ASCII DRUG_REC_ACT field and the XML <drugrecuraction> tag.

I. QUESTIONS, COMMENTS

Questions or comments may be directed to the Food and Drug Administration, Center for Drug Evaluation and Research, Office of Surveillance and Epidemiology (301-796-2360 or cderosetracking@fda.hhs.gov).