

# Medicare Fee-For-Service Data Transformation to the Sentinel Common Data Model

## User Documentation

Bradley Hammill, DrPH,<sup>1</sup> Jessica Pritchard, PhD<sup>1</sup>

*1. Department of Population Health Sciences, Duke University School of Medicine, Durham, NC*

Version 1.0

October 15, 2019

The Sentinel System is sponsored by the [U.S. Food and Drug Administration \(FDA\)](#) to proactively monitor the safety of FDA-regulated medical products and complements other existing FDA safety surveillance capabilities. The Sentinel System is one piece of FDA's [Sentinel Initiative](#), a long-term, multi-faceted effort to develop a national electronic system. Sentinel Collaborators include Data and Academic Partners that provide access to healthcare data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I.

# Medicare Fee-For-Service Data Transformation to the Sentinel Common Data Model

## User Documentation

### Table of Contents

1. Purpose .....	5
2. Disclaimers .....	5
3. Program package .....	5
3.1. Folder structure .....	5
3.1.1. Program files folder.....	5
3.1.2. Intermediate and final SCDM SAS datasets from development (testing) ETL run.....	6
3.1.3. Intermediate SCDM SAS datasets from production (final) ETL run .....	6
3.1.4. Final SCDM SAS datasets from production (final) ETL run.....	6
3.1.5. Temporary SAS datasets .....	6
3.2. SAS programs .....	6
3.2.1. 000_run_etl.sas.....	6
3.2.2. 001_etl_info.sas.....	6
3.2.3. 002_options_libs.sas.....	6
3.2.4. 003_formats.sas.....	6
3.2.5. 004_remote_macros.sas.....	6
3.2.6. 005_other_macros.sas.....	7
3.2.7. 010_etl_mbsf_enr.sas.....	7
3.2.8. 011_etl_mbsf_dem.sas.....	7
3.2.9. 012_etl_mbsf_dth.sas.....	7
3.2.10. 020_etl_pde.sas .....	7
3.2.11. 021_etl_ip.sas .....	7
3.2.12. 022_etl_snf.sas .....	7
3.2.13. 023_etl_op.sas .....	7
3.2.14. 024_etl_carr.sas.....	7
3.2.15. 030_copy_prior.sas.....	7
3.2.16. 040_combine_enr.sas.....	7
3.2.17. 041_combine_dem.sas .....	7
3.2.18. 042_combine_dth.sas.....	7
3.2.19. 050_combine_dis.sas.....	7

3.2.20.	051_combine_enc.sas.....	7
3.2.21.	052_combine_dia.sas .....	8
3.2.22.	053_combine_pro.sas.....	8
3.2.23.	060_check_etl_info.sas.....	8
3.2.24.	061_parse_log.sas.....	8
4.	Program execution.....	8
4.1.	Define ETL-specific information .....	8
4.1.1.	User input for 000_run_etl.sas .....	8
4.1.2.	User input for 001_etl_info.sas .....	9
4.1.3.	User input for 002_options_libs.sas .....	12
4.2.	Run the ETL package .....	14
5.	Output files .....	14
5.1.	Intermediate SCDM SAS datasets .....	14
5.2.	Final SCDM SAS datasets.....	14
5.3.	SAS index files .....	15
5.4.	Log and listing files.....	15

## History of Modifications

<b>Version</b>	<b>Date</b>	<b>Modification</b>	<b>Author</b>
1.0	10/15/2019	Original Version	Duke University School of Medicine

## Abbreviations

<b>Abbreviation</b>	<b>Description</b>
CARR	Carrier
CMS	Centers for Medicare & Medicaid Services
ETL	Extract, transform, and load
FFS	Fee-for-service
IP	Inpatient
MBSF	Medicare Beneficiary Summary File (formerly Denominator File)
OP	Outpatient
PDE	Part D Events (i.e., outpatient medication dispensing events)
RIF	Research identifiable file
SCDM	Sentinel Common Data Model
SNF	Skilled Nursing Facility
VRDC	Virtual Research Data Center

## 1. Purpose

This document describes the program package that enables the transformation of Medicare fee-for-service (FFS), research-identifiable files (RIFs) into the [Sentinel Common Data Model \(SCDM\) v 7.0.0 format](#). It also provides instructions on how to run the program package.

Please see the accompanying technical specifications document for detailed information about the logic and rules for transformation of these data into the SCDM. In that document, as below, the phrase “source files” refers to the original Medicare RIF data.

## 2. Disclaimers

The programs in this package are specific to the processing of the 100% Medicare data—structured using [Version K](#) with long SAS variable names—within the Center for Medicare & Medicaid Services’ (CMS) Virtual Research Data Center (VRDC). These programs will not work without modification for the transformation of Medicare data structured differently, although the programs can serve as a basis for adaptation to run in other environments.

The VRDC user accounts provisioned for the Duke Department of Population Health Sciences (DPHS) for use in the Sentinel program have extra memory, which is useful for implementing SAS hash objects to enable faster data merging on large datasets. VRDC users without additional memory will not be able to run the programs on the 100% Medicare data, although they may be able to run the programs on smaller samples of the data. Additionally, the VRDC user accounts provisioned for DPHS for use in the Sentinel program have access to data not typically made available to researchers, such as more timely, interim Part D data. VRDC users can only transform data to which they have approved access.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

## 3. Program package

### 3.1. Folder structure

There is no required folder structure for this package, but the user may wish to set up distinct folders for each purpose or type of file listed below.

#### 3.1.1. Program files folder

All SAS programs from this package must be located in a single directory. The location of this directory is set in the 000\_run\_etl.sas program.

### 3.1.2. Intermediate and final SCDM SAS datasets from development (testing) ETL run

This package allows users to test the programs on a 5% sample of the FFS Medicare data. If this option is utilized, all output SAS datasets are stored in a single directory. The location of this directory is set in the 002\_options\_libs.sas program.

### 3.1.3. Intermediate SCDM SAS datasets from production (final) ETL run

The intermediate SCDM datasets that result from the execution of this package, when not in testing or development mode, are stored in this directory. The location of this directory is set in the 002\_options\_libs.sas program.

### 3.1.4. Final SCDM SAS datasets from production (final) ETL run

The final SCDM datasets that result from the execution of this package, when not in testing or development mode, are stored in this directory. The location of this directory is set in the 002\_options\_libs.sas program.

### 3.1.5. Temporary SAS datasets

Because this package utilizes remote SAS sessions to parallel process multiple months of data simultaneously, a directory accessible by all remote sessions and by the parent session is required to store temporary datasets. Users cannot rely on SAS work directories for this purpose. This directory is not expected to store data long term and is regularly cleared by different programs in this package. As a result, this directory must be distinct from any other directory listed above. The location of this directory is set in the 002\_options\_libs.sas program.

## 3.2. SAS programs

This program package includes the following 24 SAS programs. Program outputs are described in later sections.

### 3.2.1. 000\_run\_etl.sas

This master program controls the program flow of the ETL process. This program must be edited to provide the programming root directory and to set the flags that control the ETL process. These flags are defined in Section 4.1.1. This program is the first of three programs that require editing by the user. This is the only program that is executed by the user.

### 3.2.2. 001\_etl\_info.sas

This program must be edited to provide detailed information about the years/months of source data to include in the current ETL process. The required user input is defined in Section 4.1.2. This program is the second of three programs that require editing by the user.

### 3.2.3. 002\_options\_libs.sas

This program must be edited to define the SAS library names and locations used during the ETL process and to set various options relevant to the ETL processing. These options are detailed in Section 4.1.3. This program is the third of three programs that require editing by the user.

### 3.2.4. 003\_formats.sas

Defines the SAS formats needed for the ETL process

### 3.2.5. 004\_remote\_macros.sas

Contains the macros necessary for remote submission of jobs, via SAS/GRID or locally

### 3.2.6. 005\_other\_macros.sas

Contains non-remote-submission macros common to multiple ETL programs

### 3.2.7. 010\_etl\_mbsf\_enr.sas

Transforms source Medicare Beneficiary Summary File (MBSF) enrollment data to intermediate SCDM ENROLLMENT tables. Additionally outputs the list of enrolled beneficiary identifiers for filtering source data used for other SCDM tables.

### 3.2.8. 011\_etl\_mbsf\_dem.sas

Transforms source MBSF demographics data to intermediate SCDM DEMOGRAPHIC tables

### 3.2.9. 012\_etl\_mbsf\_dth.sas

Transforms source MBSF death data to intermediate SCDM DEATH tables

### 3.2.10. 020\_etl\_pde.sas

Transforms source monthly Part D Event files (outpatient medication dispensing) to intermediate SCDM DISPENSING tables

### 3.2.11. 021\_etl\_ip.sas

Transforms source monthly FFS inpatient (IP) claims files to intermediate SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE tables

### 3.2.12. 022\_etl\_snf.sas

Transforms source monthly FFS Skilled Nursing Facility (SNF) claims files to intermediate SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE tables

### 3.2.13. 023\_etl\_op.sas

Transforms source monthly FFS outpatient (OP) claims files to intermediate SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE tables

### 3.2.14. 024\_etl\_carr.sas

Transforms source monthly FFS carrier claims (CARR) files to intermediate SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE tables

### 3.2.15. 030\_copy\_prior.sas

Copies intermediate files from the prior ETL's preliminary dataset library, if needed, as designated by the user in 001\_etl\_info.sas

### 3.2.16. 040\_combine\_enr.sas

Combines intermediate SCDM ENROLLMENT tables into a single final table

### 3.2.17. 041\_combine\_dem.sas

Combines intermediate SCDM DEMOGRAPHIC tables into a single final table

### 3.2.18. 042\_combine\_dth.sas

Combines intermediate SCDM DEATH tables into single final table

### 3.2.19. 050\_combine\_dis.sas

Combines intermediate SCDM DISPENSING tables into a single final table

### 3.2.20. 051\_combine\_enc.sas

Combines intermediate IP, OP, SNF, and CARR SCDM ENCOUNTER tables into a single final table



### 3.2.21. 052\_combine\_dia.sas

Combines intermediate IP, OP, SNF, and CARR SCDM DIAGNOSIS tables into a single final table

### 3.2.22. 053\_combine\_pro.sas

Combines intermediate IP, OP, SNF, and CARR SCDM PROCEDURE tables into a single final table

### 3.2.23. 060\_check\_etl\_info.sas

Outputs user-specific ETL information and options—from the 001\_etl\_info.sas and 002\_options\_libs.sas programs—to the listing, for purposes of checking user inputs

### 3.2.24. 061\_parse\_log.sas

Checks the ETL log for errors, warnings, and other important information, including dataset counts, timing results, etc.

## 4. Program execution

### 4.1. Define ETL-specific information

The user must define a number of macro variables and must provide specific source data information prior to executing this program package.

#### 4.1.1. User input for 000\_run\_etl.sas

The table below defines the macro variables set by the user within this program.

*Table 1. User input for 000\_run\_etl.sas*

Macro Variable	Description
ETLPROG	Path to the directory containing this package’s program files. This is the location of the “Programs files folder” described in Section 3.1.
<i>The flags below define the scope of the ETL processing</i>	
RUN_ALL	Set to Y to run the entire ETL process. If set to Y, the step-specific flags described below are ignored. Set to N to run specific parts of the ETL process. If set to N, the step-specific flags described below must be set to Y or N as desired.
ETL_MBSF_ENR	Set to Y to run the source MBSF to intermediate SCDM ENROLLMENT table transformation. Set to N to skip.
ETL_MBSF_DEM	Set to Y to run the source MBSF to intermediate SCDM DEMOGRAPHIC table transformation. Set to N to skip.
ETL_MBSF_DTH	Set to Y to run the source MBSF to intermediate SCDM DEATH table transformation. Set to N to skip.
ETL_PDE	Set to Y to run the source PDE to intermediate SCDM DISPENSING table transformation. Set to N to skip.
ETL_IP	Set to Y to run the source FFS IP to intermediate SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE table transformation Set to N to skip.
ETL_SNF	Set to Y to run the source FFS SNF to intermediate SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE table transformation Set to N to skip.

Macro Variable	Description
ETL_OP	Set to Y to run the source FFS OP to intermediate SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE table transformation Set to N to skip.
ETL_CARR	Set to Y to run the source FFS CARR to intermediate SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE table transformation Set to N to skip.
COPY_PRIOR	Set to Y to copy intermediate files from a prior ETL, if needed. Set to N to skip.
COMBINE_ENR	Set to Y to combine the intermediate SCDM ENROLLMENT files into a single final file. Set to N to skip.
COMBINE_DEM	Set to Y to combine the intermediate SCDM DEMOGRAPHIC files into a single final file. Set to N to skip.
COMBINE_DTH	Set to Y to combine the intermediate SCDM DEATH files into a single final file. Set to N to skip.
COMBINE_DIS	Set to Y to combine the intermediate SCDM DISPENSING files into a single final file. Set to N to skip.
COMBINE_ENC	Set to Y to combine the intermediate SCDM ENCOUNTER files into a single final file. Set to N to skip.
COMBINE_DIA	Set to Y to combine the intermediate SCDM DIAGNOSIS files into a single final file. Set to N to skip.
COMBINE_PRO	Set to Y to combine the intermediate SCDM PROCEDURE files into a single final file. Set to N to skip.

#### 4.1.2. User input for 001\_etl\_info.sas

This file allows users to provide detailed information about the source data that is transformed by this package. There are three different input blocks that require editing; all are structured text that is read by SAS as raw input. As such, this input must follow a `datalines` statement and be followed by a semicolon (`;`) and a `run` statement.

Note that by allowing the specification about whether to copy prior intermediate SCDM files or to (re)create them in the current process, users are able to flexibly define the ETL build type. The technical specification document describes three general build types—complete rebuild, incremental build, hybrid build—that are all supported by this functionality.

##### 4.1.2.1. Input block 1: ETL information for source MBSFs

This input block contains information about the MBSFs that are used to populate the SCDM DEMOGRAPHICS, ENROLLMENT, and DEATH files. The entire period of time from the minimum ETL date to the maximum ETL date must be covered by the lines of this input block. Each calendar year that is part of the ETL process must be represented by only one line of the input block. The specific information entered on each line is shown in the table below.

Table 2. ETL information for source MBSFs

Field Name	Description
YR	Year of MBSF (Master Beneficiary Summary File) specified below. This value is used as the file suffix for derived files.
MSTART	First month of information used for this year. The month is designated numerically. Is almost always 1 (January), unless the minimum ETL date occurs in this year and is set to something other than January.
MEND	Last month of information used for this year. The month is designated numerically. Is usually 12 (December), unless the maximum ETL date occurs in this year and is set to something other than December.
METHOD	Set to “ETL” to create the intermediate SCDM files (DEMOGRAPHICS, ENROLLMENT, DEATH) for this year from the source file. Set to “COPY” to copy the intermediate SCDM files (DEMOGRAPHICS, ENROLLMENT, DEATH) for this year from the prior ETL.
SRCDS	This is the fully-specified (library.dataset) source MBSF dataset name to use for this calendar year. Use the final annual MBSF files, where possible. For years where only quarterly data are available, use the MBSF file from the latest available quarter, as it is a cumulative file for the calendar year.

The example input block below instructs the program package to (a) copy, from the prior ETL, intermediate SCDM files derived from MBSFs for 2010 to 2016, and (b) create new intermediate SCDM files from MBSFs for 2017 (full year) and 2018 (January through September). Note that the RIF source file on the last line is the MBSF that accompanies the Q3/2018 data. As with all MBSFs, it is cumulative for the calendar year and includes information about enrollment, eligibility, etc. for the listed quarter and for all prior quarters.

```
* YR  MSTART  MEND  METHOD  SRCDS;
datalines;
2010    1      12    COPY   MBSF.MBSF_ABCD_2010
2011    1      12    COPY   MBSF.MBSF_ABCD_2011
2012    1      12    COPY   MBSF.MBSF_ABCD_2012
2013    1      12    COPY   MBSF.MBSF_ABCD_2013
2014    1      12    COPY   MBSF.MBSF_ABCD_2014
2015    1      12    COPY   MBSF.MBSF_ABCD_2015
2016    1      12    COPY   MBSF.MBSF_ABCD_2016
2017    1      12    ETL    MBSF.MBSF_ABCD_2017
2018    1      9     ETL    RIFQ2018.MBSF_ABCD_2018Q3
;
run;
```

#### 4.1.2.2. Input block 2: ETL information for source medical utilization files

This input block contains information about the FFS medical utilization files—IP, OP, CARR, and SNF claims—that populate the SCDM ENCOUNTER, DIAGNOSIS, and PROCEDURE files. The entire period of time from the minimum ETL date to the maximum ETL date must be covered by the lines of this input block. Multiple lines may appear for each calendar year, if needed. This is typically required for incremental builds, where some quarterly data are static (and intermediate SCDM files can be copied from a prior ETL) and some quarterly data for the same calendar year are new (and are transformed in the current process). The specific information entered on each line is shown in the table below.

Table 3. ETL information for source medical utilization files

Field Name	Description
YR	Year of utilization file
MSTART	First month of utilization information used. The month is designated numerically.
MEND	Last month of utilization information used. The month is designated numerically.
SUFFIX	Suffix to assign to the generated intermediate SCDM tables. This is typically just the calendar year, but can be made quarter-specific, if needed. They must be unique across all lines in this input block.
SRCLIB	The SAS libname for the source utilization files
ENRYR	The year of the intermediate SCDM Enrollment file that corresponds to the time period for this utilization period
IP	Set to "ETL" to create the intermediate SCDM files (ENCOUNTER, DIAGNOSIS, PROCEDURE) from the inpatient source files for this time period. Set to "COPY" to copy the intermediate SCDM files (ENCOUNTER, DIAGNOSIS, PROCEDURE) based on the inpatient source files for this time period from the prior ETL.
OP	Set to "ETL" to create the intermediate SCDM files (ENCOUNTER, DIAGNOSIS, PROCEDURE) from the outpatient source files for this time period. Set to "COPY" to copy the intermediate SCDM files (ENCOUNTER, DIAGNOSIS, PROCEDURE) based on the outpatient source files for this time period from the prior ETL.
CARR	Set to "ETL" to create the intermediate SCDM files (ENCOUNTER, DIAGNOSIS, PROCEDURE) from the carrier source files for this time period. Set to "COPY" to copy the intermediate SCDM files (ENCOUNTER, DIAGNOSIS, PROCEDURE) based on the carrier source files for this time period from the prior ETL.
SNF	Set to "ETL" to create the intermediate SCDM files (ENCOUNTER, DIAGNOSIS, PROCEDURE) from the SNF source files for this time period. Set to "COPY" to copy the intermediate SCDM files (ENCOUNTER, DIAGNOSIS, PROCEDURE) based on the SNF source files for this time period from the prior ETL.

The example input block below instructs the program package to (a) copy, from the prior ETL, intermediate SCDM files derived from IP, OP, CARR, and SNF files for 2010 to 2016, and (b) create new intermediate SCDM files from IP, OP, CARR, and SNF files for 2017 (full year) and 2018 (January through September).

```
* YR  MSTART  MEND  SUFFIX  SRCLIB  ENRYR  IP  OP  CARR  SNF;
datalines;
2010  1  12  2010  RIF2010  2010  COPY  COPY  COPY  COPY
2011  1  12  2011  RIF2011  2011  COPY  COPY  COPY  COPY
2012  1  12  2012  RIF2012  2012  COPY  COPY  COPY  COPY
2013  1  12  2013  RIF2013  2013  COPY  COPY  COPY  COPY
2014  1  12  2014  RIF2014  2014  COPY  COPY  COPY  COPY
2015  1  12  2015  RIF2015  2015  COPY  COPY  COPY  COPY
2016  1  12  2016  RIF2016  2016  COPY  COPY  COPY  COPY
2017  1  12  2017  RIF2017  2017  ETL  ETL  ETL  ETL
2018  1  9  2018Q  RIFQ2018  2018  ETL  ETL  ETL  ETL
;
run;
```

#### 4.1.2.3. Input block 3: ETL information for source dispensing files

This input block contains information about the outpatient prescription medication dispensing (Part D Event) files that are used to populate the SCDM Dispensing files. The entire period of time from the

minimum ETL date to the maximum ETL date must be covered by the lines of this input block. As with the utilization input bloc, described above, multiple lines may appear for each calendar year, if needed. The specific information entered on each line is shown in the table below.

*Table 4. ETL information for source dispensing files*

Field Name	Description
YR	Year of dispensing file
MSTART	First month of dispensing information used. The month is designated numerically.
MEND	Last month of dispensing information used. The month is designated numerically.
METHOD	Set to "ETL" to create the intermediate SCDM DISPENSING file from the source Part D Event files for this time period. Set to "COPY" to copy the intermediate SCDM DISPENSING file for this time period from the prior ETL.
SUFFIX	Suffix to assign to the generated intermediate SCDM tables.
SRCLIB	The SAS libname for the source dispensing files
ENRYR	The year of the intermediate SCDM Enrollment file that corresponds to the time period for this dispensing period

The example input block below instructs the program package to (a) copy, from the prior ETL, intermediate SCDM DISPENSING files for 2010 to 2016, and (b) create new intermediate SCDM DISPENSING files from source Part D Event files for 2017 (full year) and 2018 (January through September).

```

* YR  MSTART  MEND  METHOD  SUFFIX  SRCLIB  ENRYR;
datalines;
2010      1      12      COPY      2010      PDE2010      2010
2011      1      12      COPY      2011      PDE2011      2011
2012      1      12      COPY      2012      PDE2012      2012
2013      1      12      COPY      2013      PDE2013      2013
2014      1      12      COPY      2014      PDE2014      2014
2015      1      12      COPY      2015      PDE2015      2015
2016      1      12      COPY      2016      PDE2016      2016
2017      1      12      ETL       2017      PDE2017      2017
2018      1      9       ETL       2018U     PDE2018      2018
;
run;

```

#### 4.1.3. User input for 002\_options\_libs.sas

The table below defines the macro variables set by the user within this program. Some pre-specified, general SAS options can also be found, and modified, at the end of this program.

Table 5. User input for 002\_options\_libs.sas

Macro Variable	Description
<i>The variables below define ETL-specific information.</i>	
ETLNUM	Number of the current ETL. In the distributed code, this information is used to define the location of multiple SAS library locations. This macro variable is not required if the library specifications are changed and no longer depend on this information.
ETLVRSN	Version of the current numbered ETL. In the distributed code, this information is used to define the location of multiple SAS library locations. This macro variable is not required if the library specifications are changed and no longer depend on this information.
MINDT	The minimum date for encounters and other information in the current ETL process. Should be formatted as "DDMONYYYY"d (e.g., "01JAN2010"d).
MAXDT	The maximum date for encounters and other information in the current ETL process. Should be formatted as "DDMONYYYY"d (e.g., "31DEC2018"d).
LASTNUM	Number of the latest prior production ETL. In the distributed code, this information is used to define the location of multiple SAS library locations. This macro variable is not required if the library specifications are changed and no longer depend on this information.
LASTVRSN	Number of the latest prior production numbered ETL. In the distributed code, this information is used to define the location of multiple SAS library locations. This macro variable is not required if the library specifications are changed and no longer depend on this information.
ETLDEV	Set to Y to run the ETL process on a 5% sample of FFS Medicare beneficiaries for testing/development. Set to N for the production run of the ETL process, using all data available.
<i>The variables below define remote SAS session processing information.</i>	
NUMSESS	Number of remote sessions available to open for parallel processing.
USEGRID	Set to Y if SAS/GRID is enabled in the computing environment and must be used for remote session management. Set to N to use the local machine for remote sessions.
GRIDSRV	The name of the SAS/GRID server used for remote sessions; required if USEGRID = Y
<i>The SAS LIBNAMEs below specify the locations of various SAS datasets</i>	
SAS LIBNAMEs for source data in the VRDC are preset for SAS Enterprise Guide. If additional SAS LIBNAMEs are needed, enter them in this file.	
FINAL	Path to the directory in which to store the final SCDM datasets from the production run for the current ETL
DEV	Path to the directory in which to store the intermediate and final SCDM datasets from the development run for the current ETL
PRELIM	Path to the directory in which to store the intermediate SCDM datasets from the production run for the current ETL
TEMP	Path to the directory in which to store temporary files for jobs submitted to remote sessions
LASTPROD	Path to the directory in which the final SCDM datasets from the production run for the latest prior ETL are stored
LASTDEV	Path to the directory in which the intermediate and final SCDM datasets from the development run for the latest prior ETL are stored
LASTPREL	Path to the directory in which the intermediate SCDM datasets from the production run for the latest prior ETL are stored

## 4.2. Run the ETL package

Run the 000\_run\_etl.sas program in batch mode. This is the only program that requires execution. It incorporates other programs as needed.

## 5. Output files

The following files are generated by this ETL package.

### 5.1. Intermediate SCDM SAS datasets

Multiple files of each type listed below are generated for any specific year [yyyy] or time period [suffix], according to the user input in 001\_etl\_info.sas. These have a file extension of .sas7bdat.

*Table 6. Intermediate SCDM SAS datasets*

Dataset Name	Description
DEATH_[yyyy]	Intermediate SCDM DEATH file
DEMOGRAPHIC_[yyyy]	Intermediate SCDM DEMOGRAPHIC file
DISPENSING_[suffix]	Intermediate SCDM DISPENSING file
DX_CARR_[suffix]	Intermediate SCDM DIAGNOSIS file, from Carrier claims
DX_IP_[suffix]	Intermediate SCDM DIAGNOSIS file, from Inpatient claims
DX_OP_[suffix]	Intermediate SCDM DIAGNOSIS file, from Outpatient claims
DX_SNF_[suffix]	Intermediate SCDM DIAGNOSIS file, from SNF claims
ENC_CARR_[suffix]	Intermediate SCDM ENCOUNTER file, from Carrier claims
ENC_IP_[suffix]	Intermediate SCDM ENCOUNTER file, from Inpatient claims
ENC_OP_[suffix]	Intermediate SCDM ENCOUNTER file, from Outpatient claims
ENC_SNF_[suffix]	Intermediate SCDM ENCOUNTER file, from SNF claims
ENROLLMENT_[yyyy]	Intermediate SCDM ENROLLMENT file
IDS_ENROLL_[yyyy]	List of Enrolled IDs
PX_CARR_[suffix]	Intermediate SCDM PROCEDURE file, from Carrier claims
PX_IP_[suffix]	Intermediate SCDM PROCEDURE file, from Inpatient claims
PX_OP_[suffix]	Intermediate SCDM PROCEDURE file, from Outpatient claims
PX_SNF_[suffix]	Intermediate SCDM PROCEDURE file, from SNF claims

### 5.2. Final SCDM SAS datasets

These have a file extension of .sas7bdat.

*Table 7. Final SCDM SAS datasets*

Dataset Name	Description
DEATH	Final SCDM DEATH file
DEMOGRAPHIC	Final SCDM DEMOGRAPHIC file
DIAGNOSIS	Final SCDM DIAGNOSIS file
DISPENSING	Final SCDM DISPENSING file
ENCOUNTER	Final SCDM ENCOUNTER file
ENROLLMENT	Final SCDM ENROLLMENT file
PROCEDURE	Final SCDM PROCEDURE file



### 5.3. SAS index files

Indexes are created for many final SCDM SAS datasets. These have a file extension of .sas7bndx and are stored in the same directory as the final SCDM SAS datasets.

*Table 8. SAS index files*

Index Name	Description
DEATH	Final SCDM DEATH file index
DEMOGRAPHIC	Final SCDM DEMOGRAPHIC file index
DISPENSING	Final SCDM DISPENSING file index
ENROLLMENT	Final SCDM ENROLLMENT file index

### 5.4. Log and listing files

Log and listing files are run-specific. In the file names below, [yyyymmdd] represents the date that 000\_run\_etl.sas was submitted.

*Table 9. Log and listing files*

File Name	Description
_etl_dev_[yyyymmdd].log	Log file for development run.
_etl_dev_[yyyymmdd].lst	Listing file for development run.
_etl_prod_[yyyymmdd].log	Log file for production run.
_etl_prod_[yyyymmdd].lst	Listing file for production run.