Chronic Condition Data Warehouse

Your source for national CMS Medicare and Medicaid research data

CCW User Guide:
T-MSIS Analytic Files (TAF)
Research Identifiable Files (RIFs)

November 2019

Version 1.0

This page intentionally left blank.

Table of Contents

List of Tables

List of Figures

List of Data Tips

Overview

The Centers for Medicare & Medicaid Services (CMS) uses the Chronic Conditions Warehouse (CCW) to develop and manage CMS research data sources. Research Identifiable Files, or RIFs, are disseminated to academic researchers and certain government agencies which have been approved under a Data Use Agreement (DUA) to obtain CMS data for research purposes. CMS has authorized the release of a new format of Medicaid and CHIP data to the research community called the Transformed Medicaid Statistical Information System (T-MSIS). This user guide describes the new T-MSIS research files, their strengths and weaknesses compared to previous Medicaid data formats, and how to obtain them.

Medicaid is a state-administered benefit with guidance/requirements and shared funding from the federal government and individual state governments (aka Title XIX of the Social Security Act from 1965). Each state must provide the minimum federally mandated services and coverage for federally mandated eligibility groups; however, state benefits may and do vary from state to state. Many groups of people are covered by Medicaid, depending on the state's requirements (e.g., age, pregnancy status, disabled, blind, or aged, income level and resources, U.S. citizenship or lawful immigration status). There are also special rules for those who live in nursing homes and for children with disabilities living at home. CMS administers the Medicaid program jointly with states. Please see the CMS Medicaid.gov website for additional information regarding Medicaid (https://www.medicaid.gov/).

The Balanced Budget Act of 1997 amended the Social Security Act to add Title XXI, which is the State Children's Health Insurance Program (aka SCHIP which has evolved to simply be called CHIP). CHIP enables states to extend insurance coverage to low-income children who are not eligible for Medicaid and who do not have private insurance. States have the flexibility to design their own program within Federal guidelines. Benefits vary by state and the type of CHIP program (i.e., whether the CHIP program operates as an expansion of Medicaid [sometimes referred to as M-CHIP], a separate child health insurance program [referred to as S-CHIP], or a combination of the two). For more information on the CHIP program, see the CMS website https://www.medicaid.gov/chip/index.html).

The Patient Protection and Affordable Care Act of 2010 (ACA) expanded insurance coverage through offering private individual policies via a centralized "Marketplace" and through state Medicaid expansions. The ACA expands Medicaid coverage for most low-income adults to 138% of the federal poverty level (FPL). As of 2014, states may decide whether to adopt Medicaid expansion; not all states have expanded their Medicaid programs.[1]

Each state compiles information regarding their Medicaid and CHIP enrollment, service utilization, and payments in the recently implemented T-MSIS format, and provides T-MSIS data files to CMS. Using the T-MSIS data files from states, CMS creates the T-MSIS Analytic Files (TAFs). The CCW obtains the TAF files, loads them to a database and creates claims and enrollment Research Identifiable Files (RIFs). These files can be disseminated to academic researchers and certain government agencies with an approved DUA for research purposes. The TAF RIFs contain personally identifiable information (PII), and are subject to the Privacy Act and other Federal government rules and regulations (see the ResDAC web site, https://www.resdac.org/ for information on requesting TAF RIF data).

[1] HealthCare.gov. Medicaid & CHIP. "Medicaid expansion & what it means for you." https://www.healthcare.gov/medicaid-chip/medicaid-expansion-and-you// Accessed 09/25/2019

For Medicaid data from 1999 through 2014[2], CMS produced annual MAX (Medicaid Analytic eXtract) files, which were created using data derived from the Medicaid Statistical Information System (MSIS). MSIS, and therefore MAX, has been phased out as states converted their Medicaid and CHIP data submissions to CMS into the new T-MSIS format.

T-MSIS represents the next generation of national data for Medicaid and CHIP beneficiaries and the services they use. T-MSIS differs from MSIS in a number of important ways, including the timing of submissions (monthly vs. quarterly) and the amount of content states report (nearly four times as many data elements, including several new segments).

The TAF research files are linked by a unique CCW person identifier (the CCW beneficiary identifier; variable called the BENE_ID), allowing researchers to analyze information across the entire continuum of care, both within and across years, without using standard person identifiers such as MSIS_ID, Social Security number, or Medicare Health Insurance Claim number for dual eligible beneficiaries. This unique CCW identifier follows an enrollee across states and other CCW research data sources, so that if the same person is enrolled in Medicaid or CHIP in more than one state or is also a Medicare beneficiary, those records may be combined using the BENE_ID. For example, the CCW also contains Medicare enrollment data, fee-for-service (FFS) claims data, Medicare Advantage encounter data submitted by plans[3] to CMS, and assessment data (e.g., Minimum Data Set [MDS] and Outcome and Assessment Information Set [OASIS]). The BENE_ID facilitates analysis across all CMS data sources in the CCW.

There are five types of data files available for TAF RIF at this time, including the following:

1) Annual Demographic and Eligibility (DE),
2) Inpatient (IP) Claims,
3) Long Term Care (LT) Claims,
4) Pharmacy (RX) Claims, and
5) Other Services (OT) Claims

This guide provides users with information that may be helpful in understanding and requesting the TAF research files. For example, a variety of resources are available to assist TAF data users, as described in Chapter 6. Further Assistance with Medicaid Data. Record layouts for the TAF research files are available on the ccwdata.org website (https://www.ccwdata.org/web/guest/data-dictionaries). And a list of acronyms used throughout the paper is included as Appendix A - List of Acronyms. Related to acronyms, the TAF research files, or TAF RIFs, are referred to as "TAFs" throughout the rest of this user guide for brevity.

[2] Not all states have MAX files for 2014; 2013 MAX may be the last year some states are included with this data format. Refer to -Appendix B – State Cutover to T-MSIS and Availability of MAX and TAF Research Files

[3] Medicare Advantage Encounter data files are available starting with 2015.

Chapter 1. Introduction

The Centers for Medicare & Medicaid Services (CMS) utilize the Chronic Conditions Warehouse (CCW) to develop and manage CMS research data resources. TAF RIFs are created from enrollment and service-related data submitted by state Medicaid and CHIP programs to CMS. Medicaid and CHIP service data sent by states are typically obtained from either FFS claims or managed care encounter records. Approximately 75 million people were enrolled in Medicaid or CHIP in 2016.[4] Accordingly, the associated enrollment and claims data files are very large.

The CCW Medicare, Medicaid, and CHIP data are subject to the HIPAA Privacy Rule, the Privacy Act and other federal government rules and regulations. Researchers need to obtain a Data Use Agreement (DUA) to obtain the files. Additional details regarding the data request process can be found in Chapter 5. Receiving CCW Data.

A. Source Data

The administration of Medicaid and CHIP programs is a state-federal partnership. Medicaid and CHIP are administered by states, according to federal requirements. States are responsible for eligibility verification and enrollment, and providing benefits to qualifying recipients within their state[5].

CMS requires that states share key information regarding the Medicaid and CHIP programs. Since 1999, states have been required to submit electronic Medicaid claims and eligibility data files through the Medicaid Statistical Information System, or MSIS. States began planning for and transitioning to the Transformed Medicaid Statistical Information System (T-MSIS) as early as 2013. However, initial T-MSIS reporting dates (sometimes referred to as "cutover dates") vary by state.

Each state selected a starting calendar quarter for T-MSIS data, which corresponded to the last quarter for which MSIS data was submitted. All states were required to cut over to the T-MSIS format by October 2015, although some began submitting in this format earlier. State-specific cutover dates are listed in Appendix B – State Cutover to T-MSIS and Availability of MAX and TAF Research Files.

TAF RIF data are available for each state starting with the first full calendar year after they began submitting T-MSIS data (Table 1)[6]. By 2016, all states[7] had cutover to T-MSIS. However, Arkansas and Puerto Rico were not included in the initial 2015-2016 TAF RIF release because they are missing information for key variable segments, such as enrollment.

---

[4] US HHS, CMS Monthly Medicaid Enrollment Reportshttps://www.medicaid.gov/medicaid/program-information/downloads/updated-december-2016-enrollment-data.pdfhttps://www.medicaid.gov/medicaid/program-information/downloads/updated-december-2016-enrollment-data.pdf Accessed 10/22/2019.

[5] There is additional information regarding the state-federal partnership, and state responsibilities for data collection and submission to CMS on the Medicaid.gov website (e.g., https://www.medicaid.gov/state-overviews/scorecard/national-context/index.html). Accessed 09/25/2019.

[6] Iowa and Pennsylvania began submitting CHIP data prior to Medicaid data, but are not included in the TAF until their Medicaid cutover date.

[7] The Virgin Islands, Montana Third-Party administrative claims, and Wyoming CHIP are not yet submitting to T-MSIS and are therefore not included in the 2014-2016 TAF RIF.

Table 1. Count of States with TAFs in CCW, by Year

	2014	2015	2016
Number of States	19	30	50*

* All states and DC except for Arkansas, Puerto Rico, Montana Third-Party Administrative claims, Wyoming CHIP, and the Virgin Islands.

States that did not submit T-MSIS data for all 12 months of the year are not included in the RIFs for that service year (or earlier service years). For these states, MAX data files should be requested[8]. The MAX files are discussed further in the next section.

B. Changes from MAX Data

CMS produced the MAX files from state-submitted MSIS data from 1999 to 2013. MAX data contains five types of files which correspond to the initial release of T-MSIS RIF file types:

1) Person Summary File (PS) – similar to the Annual Demographic & Eligibility File in TAF
2) Hospital Inpatient File (IP)
3) Long Term Care File (LT)
4) Prescription Drug File (RX)
5) Other Services File (OT)

MAX was phased out as states converted their data submissions to the T-MSIS format. Depending on the state(s) of interest, for 2014 and 2015, researchers may need to request MAX files for some states and TAFs for other states. See Appendix B – State Cutover to T-MSIS and Availability of MAX and TAF Research Files to determine the type of data file available for each state and service year.

The new TAFs contain information on a broader set of individuals than MAX, and includes Medicaid FFS and managed-care enrolled, M-CHIP or Separate CHIP (also referred to as S-CHIP) enrolled, and Medicaid expansion populations, when applicable. MAX does not include CHIP claims. Note that throughout this document, we use the acronym CHIP to mean both M-CHIP and S-CHIP, unless otherwise specified. There are also more than twice as many data elements submitted by states through T-MSIS, compared to MSIS.

For example:

- New demographic information is available (e.g., household size),
- Additional program enrollment/benefit information is present (e.g., more waiver details, and more managed care enrollment details), and
- 2-part claims: header – plus all claim line records (MAX limited claim line information to 26 lines)

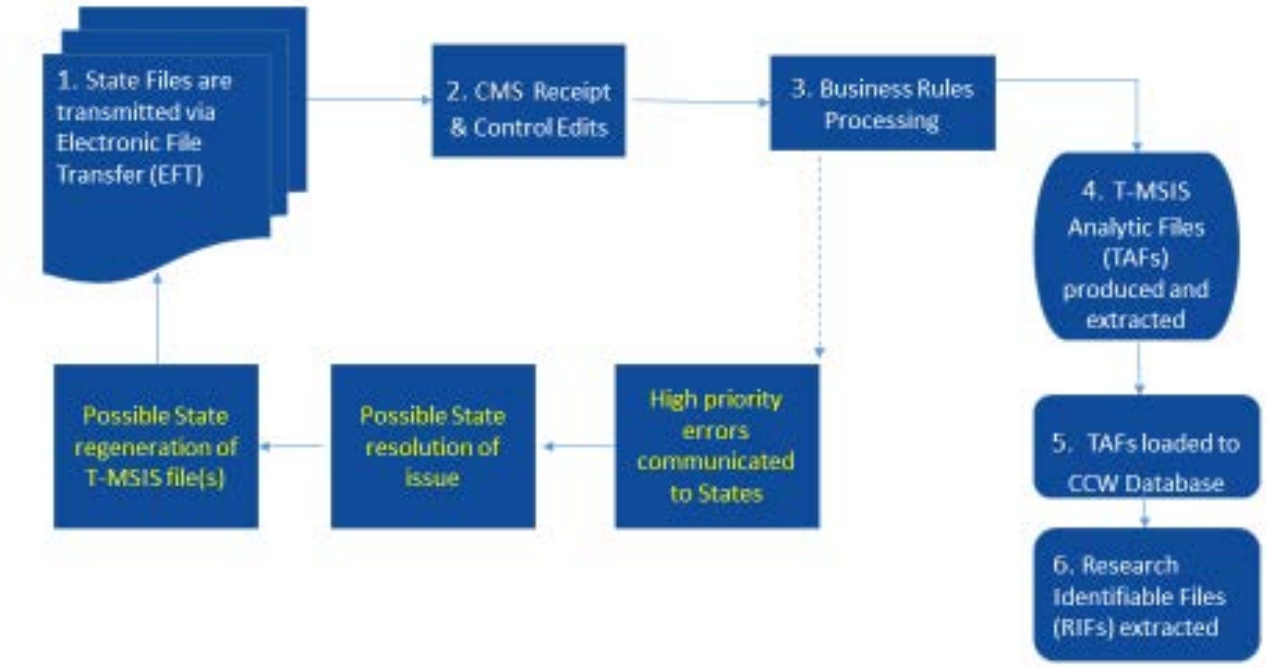
[8] For states that do not have a full 12 months of T-MSIS data, CMS produces MAX files. See https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAXGeneralInformation.html Accessed 09/25/2019.

C. Creation of TAF Research Files

This section of the document contains information about the Medicaid and CHIP source data used to create the TAF research files. States collect and process enrollment, claims, managed care plan and provider data and submit that data to CMS via T-MSIS. CMS creates standard T-MSIS Analytic Files (TAFs) and sends them to the CCW, which loads them to the relational CCW database and then extracts standard RIFs to disseminate to data users.

A visual depiction of the Medicaid and CHIP data flow is in Figure 1.

Figure 1. T-MSIS Data Processing Flow



States submit transactional T-MSIS files to CMS each month. Data may be re-submitted or re-sent for all service dates. A CMS database stores each claim record submitted by states and uses an algorithm to group all records into a unique service "claims family". An original claim and any claims that adjust the original claim make up a claim family, which can consist of one or multiple claims depending on the number of adjustments. Within a family, any claim flagged as final action means that it is the final record for the service reported on all claims in the family. The TAF include only the final action claims to help make the data more manageable for analytic purposes. Additional details regarding TAF Final Action, among a variety of other key research topics, are available in a series of informational white papers called "Data Quality briefs"[9]. Data Quality (DQ) briefs are available to users on the ResDAC website at www.resdac.org.

[9] For the Data Quality brief on final action claims, see TAF DQ Brief #3011, "Final Action Status in T-MSIS Claims in 2016."

States also submit monthly enrollment files to CMS. These are accumulated, and then CMS produces the TAF annual demographic and eligibility (DE) file, which represents every beneficiary enrolled in a Medicaid or CHIP program anytime during the service year, as well as the four TAF claims files (Inpatient, Long Term Care, Other Services, and Pharmacy). The TAF files are offloaded (extracted) and sent to the CCW. CCW assigns a BENE_ID and partitions the data into calendar year RIFs based on service (or enrollment) date. See Section D, Assignment of a Beneficiary Identifier, below for more information on the beneficiary identifier. The 2014 - 2016 RIFs are created from data the CCW received in the summer of 2019.

Although claim-level payment information is present for all FFS claims, the Center for Medicaid and CHIP Services (CMCS) at CMS has required redaction of some payment fields for managed care claims due to the proprietary and confidential nature of this information. In addition, some state files include "service tracking claims" that are lump-sum payments to providers or plans. Service tracking records (i.e., where the variable called CLM_TYPE_CD=4, D, or X) are not included in the RIFs because they are not claims for services to a specific enrollee and are not discrete services provided to patients with corresponding diagnosis or procedure information. Additional details regarding the contents of the claims files is in Chapter 3, Section B, Content and Description of Claims RIFs.

The record layouts for all TAF research files may be obtained from the ccwdata.org website, under the Data Dictionaries tab (https://www.ccwdata.org/web/guest/data-dictionaries). Throughout this document, when we identify a particular data variable by name we will often identify the specific SAS® name by use of all capitals.

D. Assignment of a Beneficiary Identifier

The CCW beneficiary identifier (BENE_ID) is assigned to a unique person across distinct CCW data sources, thereby allowing researchers to track all services for a beneficiary over time. Further, Medicaid enrollees who are also enrolled in the Medicare program (or "dual eligible" beneficiaries) are assigned the same BENE_ID for both Medicaid and Medicare records within the CCW.

To construct the BENE_ID, the CCW team developed internal cross-reference files consisting of historical Medicaid and Medicare enrollment information using CMS data sources such as the Enterprise Cross Reference (ECR) file. When a new TAF enrollment file is received, the CCW examines the MSIS_ID, STATE_CD, social security number (SSN), date of birth, sex and other beneficiary identifying information to determine if a BENE_ID already exists.

If there is a single record that "best matches" the information in the Medicaid enrollment record, then the BENE_ID on that historical record is assigned to the Medicaid enrollment record. If there is no match or no "best match" after CCW has exhausted a stringent matching process, a null (or missing) BENE_ID is assigned to the Medicaid enrollment record.

For any given year, approximately 3.5% of TAF records have a null (or missing) BENE_ID. Once a BENE_ID is assigned to a Medicaid or CHIP enrollment record for a particular year (with the exception of those assigned to a null value), it will not change. When a new Medicaid or CHIP enrollment file is received, CCW attempts to reassign those with missing BENE_IDs. The TAF research files do include the TAF DE records that are missing a BENE_ID.

The BENE_ID from the Medicaid and CHIP enrollment record is also assigned to each respective claim record (IP, LT, RX, OT). As with enrollment records, once a BENE_ID is assigned in claims data files, it does not change. Additional guidance regarding use of BENE_ID or MSIS_ID for linking TAFs together or for aggregating data appear later in this user guide (see Chapter 4, Section A, Linkage Variables and Joining Files and Summarizing Data).

The unique BENE_ID field is specific to the CCW and is not applicable to any other identification system or data source. This identifier is encrypted prior to delivering the data files to researchers. In addition, all data files delivered to researchers are encrypted (see Encryption Information in Chapter 5. Receiving CCW Data for details). It is important to note that each research request employs a different encryption key for the beneficiary identifier field and the data files.

Chapter 2. Annual Demographic and Eligibility RIF

The Annual Demographic and Eligibility (DE) TAF RIF is an annual file that CMS creates from the monthly state enrollment data. Additional details were presented earlier, in Chapter 1, Section C, Creation of TAF Research Files. Each annual DE file includes all Medicaid and CHIP enrollees documented as being enrolled in a Medicaid or CHIP program for at least one day of the year.

Upon enrollment, all Medicaid and CHIP beneficiaries are assigned an MSIS identifier, called an MSIS ID, by the state. Note that there are some states that assign more than one MSIS ID to the same beneficiary. There is also one state, Iowa, that assigned *different* MSIS IDs for individuals enrolled in its CHIP program versus its Medicaid program.

State-assigned MSIS IDs are not ideal for research purposes because beneficiaries may have more than one MSIS ID number over time due to enrollment in Medicaid and CHIP in the same or different states. The BENE_ID allows linkage of an individual's data both within and across states, years, and different TAF RIF file types, as well as across other CCW data sources. Some beneficiaries will not have a BENE_ID; additional details are in Chapter 4, Section A, Linkage Variables and Joining Files.

The DE RIF contains several files – the "Base" or core enrollment/demographic file, and six (6) supplemental files (1) Eligibility Dates, (2) Managed Care Enrollment, (3) Waiver Program Enrollment, (4) Money Follows the Person (MFP), (5) Health Home and State Plan Options (SPO), and (6) Disability and Need. These files are described below.

A. DE Base File

The DE RIF Base is the primary eligibility file which nearly all researchers will use. It contains 188 variables with information on beneficiaries' demographic characteristics and details of their enrollment in Medicaid or CHIP. The data file has a record for each MSIS_ID within a state (STATE_CD). The BENE_ID is also populated when possible. Since the same BENE_ID may apply to more than one MSIS_ID and STATE_CD, data users should link the DE Base with the DE supplemental files using the MSIS_ID and STATE_CD to ensure the correct DE supplemental file records are associated with the selected DE Base record. Additional details are presented in Chapter 4, Section A Linkage Variables and Joining Files, and an example is shared in Table 17.

The DE Base file includes variables related to demographics and geography (see Table 2). It also includes monthly and annual type of enrollment information; how the beneficiary qualified for Medicaid or CHIP; information regarding eligibility group codes, enrollment in managed care, and indicators for enrollment in waivers (and/or Money Follows the Person demonstration, Health home, etc.). Refer to Table 3 for examples of monthly and yearly enrollment variables.

The DE Base file contains a small number of "dummy" records to represent beneficiaries where there are claims data but no eligibility record in the source T-MSIS data from the states. Since most analyses with T-MSIS will use information from the enrollment files, CMS created a record in the DE Base file for the MSIS_ID where only the MSIS_ID and STATE_CD are populated with valid values. The dummy enrollment records are identified by a missing eligibility data indicator (MISG_ELGBLTY_DATA_IND) value of 1.

Table 2. DE Base File – Examples of Key Demographic and Geographic Variables

DE Base Variables	Description	SAS variable name
Demographic Information	Date of Birth	BIRTH_DT
	Age (in years)	AGE
	Date of Death	DEATH_DT
	Sex (Biological) - Latest in Year	SEX_CD
Geographic Information	State FIPS Code for Beneficiary Home or Mailing Address - Latest in Year	BENE_STATE_CD
	County Code for Beneficiary Home or Mailing Address - Latest in Year	BENE_CNTY_CD
	ZIP Code for Beneficiary Home or Mailing Address - Latest in Year	BENE_ZIP_CD

Table 3. DE Base File – Examples of Key Enrollment Variables

DE Base Variables	Description	SAS variable name
Monthly Information	Medicaid Enrollment Days – (one variable for each month)	MDCD_ENRLMT_DAYS_01 -12
	CHIP Enrollment Days - (one variable for each month)	CHIP_ENRLMT_DAYS_01 - 12
	Medicaid, Medicaid Expansion CHIP, or Separate CHIP Code - (one variable for each month)	CHIP_CD_01 -12
	Scope of Medicaid or CHIP Benefits - (one variable for each month)	RSTRCTD_BNFTS_CD_01 -12
	Medicare-Medicaid Dual Eligibility Code - (one variable for each month)	DUAL_ELGBL_CD_01 -12
	Eligibility Group Code - (one variable for each month)	ELGBLTY_GRP_CD_01 -12
Yearly Information	Medicaid Enrollment Days - Total in Year	MDCD_ENRLMT_DAYS_YR
	CHIP Enrollment Days - Total in Year	CHIP_ENRLMT_DAYS_YR

Identifying the type and level of benefits for enrollees is central to understanding the population for whom the researcher will examine the claims. Some important subgroups that are of interest to many data users include: (1) Medicaid (Title XIX) enrolled, and (2) M-CHIP (Title XXI) enrolled, and (3) S-CHIP (Title XXI) enrolled. We highlight an important data consideration in Data Tip 1, below.

Data Tip 1. How to Identify Medicaid or CHIP Enrollment

The monthly Medicaid, Medicaid Expansion CHIP, or Separate CHIP Code field is one of the key fields to use for identifying beneficiaries enrolled in Medicaid or CHIP for the month (CHIP_CD_01 -12). Possible monthly values are:

1 = Medicaid-enrolled month
2= Medicaid Expansion CHIP (M-CHIP) month
3= Separate Title XXI CHIP (S-CHIP) enrolled month
4 = Individual was both Medicaid-Eligible and Separate CHIP eligible during the same month
Null/missing = not enrolled during the month

The CHIP_CD, along with the eligibility group code (ELGBLTY_GRP_CD_01 -12) can provide information regarding enrollment.
A SAS® coding example illustrates one method for identifying Medicaid or CHIP enrollees; of course you may adapt this code to use any statistical software you prefer:

1) Begin by using the DE RIF, which we have called “DE_BASE” in this example.
2) Create a variable called “group”, where the values are Medicaid, M-CHIP, S-CHIP, and Unclassified.
3) Use the latest CHIP_CD (CHIP_CD_LTST) and the latest eligibility group code (ELGBLTY_GRP_CD_LTST) to parse the data. For S-CHIP we also check to see if there were any CHIP enrollment days in the year (using CHIP_ENRLMT_DAYS_YR).
4) Save a file with your classifications – we call this “population”.

```
data population;
   length group $20.;
   set DE_BASE;
   if CHIP_CD_LTST = '1' and ELGBLTY_GRP_CD_LTST in ('01' '02' '03' '04' '05' '06' '07' '08' '09' '10' '11'
       '12' '13' '14' '15' '16' '17' '18' '19' '20' '21' '22' '23' '24' '25' '26' '27' '28' '29' '30' '31' '32' '33' '34'
       '35' '36' '37' '38' '39' '40''41' '42' '43' '44' '45' '46' '47' '48' '49' '50' '51' '52' '53' '54' '55' '56' '57'
       '58' '59' '60'
        '69' '70' '71' '72' '73' '74' '75') then group = "Medicaid";
          else if CHIP_CD_LTST = '2' then group = "M-CHIP";
          else if CHIP_CD_LTST = '3' and CHIP_ENRLMT_DAYS_YR > 0 then group = "S-CHIP";
          else group = "Unclassified";
run;
```

The restricted benefits code (RSTRCTD_BNFTS_CD_01 -12) also provides information regarding the scope of Medicaid or CHIP benefits (e.g., full or partial benefits).

There is state-level variability in data quality related to identification of Medicaid/CHIP enrollment, which may necessitate a variation on this logic to appropriately identify the population of interest for your study. To learn more about identifying particular populations, please refer to the DQ briefs and the technical documentation available on the ResDAC website.

In addition to Medicaid/CHIP enrollment variables mentioned above, there are fields to identify state-specific eligibility groups (variables such as STATE_SPEC_ELGBLTY_GRP_CD_01-12). There are also some

variables where an annual value is provided. These include variables that are populated with the latest valid value observed in the source data (naming convention is *_LTST), and other variables that indicate whether a program or service ever applied during the year (naming convention is *_IND in conjunction with a SAS® label containing the phrase "Ever in Calendar Year"). Refer to Table 4 for examples. Data users should refer to the TAF Enrollment Codebook to obtain additional details, including the valid values.

Table 4. DE Base File – Examples of Key Annual Enrollment Variables

DE Base Variables	Description	SAS variable name
Latest in Year	Medicaid, Medicaid Expansion CHIP, or Separate CHIP Code - Latest in Year	CHIP_CD_LTST
	Scope of Medicaid or CHIP Benefits - Latest in Year	RSTRCTD_BNFTS_CD_LTST
	Medicare-Medicaid Dual Eligibility Code - Latest in Year	DUAL_ELGBL_CD_LTST
	Eligibility Group Code - Latest in Year	ELGBLTY_GRP_CD_LTST
Ever in Year	Disability Ind - Difficulty Concentrating - Ever in Calendar Year	DSBLTY_DFCLTY_CNCNTRTNG_IND
	Disability Ind - Difficulty Walking - Ever in Calendar Year	DSBLTY_DFCLTY_WLKG_IND
	Disability Ind - Difficulty Dressing or Bathing - Ever in Calendar Year	DSBLTY_DFCLTY_DRSNG_BATHNG_IND

Not all enrollees in the DE Base file will be found in the supplemental files. To determine if a beneficiary has a record in a supplemental file, refer to the indicator variables within the DE Base. See Table 5. These nine fields are binary indicators, where 1 = beneficiary has a record in the supplemental file.

Table 5. DE Base File – Variables Indicating Records in DE Supplemental File

Description	SAS variable name
Beneficiary Record In Supplemental Dates File	SPLMTL_DTS
Beneficiary Record in Supplemental Managed Care File	SPLMTL_MC
Beneficiary Record in Supplemental Waiver File	SPLMTL_WVR
Beneficiary Record in Supplemental Health Home and State Plan Option (SPO) File	SPLMTL_HLTH_HOME_SPO
Beneficiary Record in Supplemental Money Follows Person (MFP) File	SPLMTL_MFP
Beneficiary Home and Community-Based Services (HCBS) Record in Supplemental Disability File*	SPLMTL_DSB_HCBS
Beneficiary Long-Term Services and Supports (LTSS) Record in Supplemental Disability File*	SPLMTL_DSB_LTSS
Beneficiary Lock-In Record in Supplemental Disability File*	SPLMTL_DSB_LCKIN
Beneficiary Other Needs Record in Supplemental Disability File*	SPLMTL_DSB_OTHR

* Any of these four indicators means there is a record in the Disability and Need supplemental file.

B. DE Supplemental Files

The six DE supplemental files may be used as needed to augment information in the DE Base RIF (Table 6). A description of the variables contained in the DE file can be found on the Data Dictionary page on the CCW website at http://www.ccwdata.org/web/guest/data-dictionaries.

Table 6. DE Supplemental Files

DE Supplemental File	# of variables in RIF
1. Eligibility Dates	8
2. Managed Care Enrollment	410
3. Waiver Program Enrollment	268
4. Money Follows Person (MFP)	23
5. Health Home and State Plan Options (SPO)	99
6. Disability and Need	220

Each of these six files are described briefly below.

1. Eligibility Dates

This file may have more than one record per beneficiary; that is, there may be multiple records for an MSIS_ID (within a STATE_CD) for a calendar year. Each record represents an enrollment spell for a Medicaid or CHIP beneficiary. A person may have more than one record in this supplemental file if there was more than one enrollment type code during the year (e.g., person was CHIP for part of the year and Medicaid for part of the year) or more than one enrollment spell within the year (e.g., covered for Medicaid a few months early in the year, not covered for several months, and then covered again at end of the year).

There is an enrollment start and end date (ENRLMT_START_DT and ENRLMT_END_DT) for each record. Note that enrollment days across Medicaid and CHIP may overlap. Additional information is available in DQ brief #4091, "Assessing the Validity of Effective Eligibility Dates in 2016."

2. Managed Care Enrollment

The Managed Care Enrollment supplemental file contains one record for each beneficiary (identified by each unique combination of MSIS_ID and STATE_CD) who was enrolled in some form of managed care in any month during the calendar year. The file accommodates up to 16 managed care plan type codes and associated plan IDs for each calendar month; these arrays account for 384 of the 410 variables in the file (93.6%). It also contains constructed variables that summarize the count of months a beneficiary was enrolled in each managed care plan type within the year.

3. Waiver Program Enrollment

The Waiver Program Enrollment supplemental file contains one record for each beneficiary (identified by each unique combination of MSIS_ID and STATE_CD) who was enrolled in at least one waiver program in any month during the calendar year. The file accommodates up to 10 waiver type codes and associated waiver IDs for each month; these arrays account for 240 of the 268 variables in the file (89.5%). It also contains constructed variables that summarize the count of months a beneficiary was enrolled in each type of waiver program within the year (Table 7).

Table 7. DE Waiver Program Enrollment Supplemental File – Annual Summary Variables

Description	SAS variable name
1915(c) Waiver Months	WVR_1915C_MOS
1915(b) Waiver Months	WVR_1915B_MOS
1915(b)(c) Waiver Months	WVR_1915BC_MOS
1115 Pharmacy Waiver Months	PHRMCY_WVR_1115_MOS
1115 Disaster-Related Waiver Months	DSTR_RLTD_WVR_1115_MOS
1115 Family Planning Only Waiver Months	FMLY_PLNG_ONLY_WVR_1115_MOS
1115 Health Insurance Flexibility and Accountability Waiver Months	HIFA_WVR_1115_MOS
1115 Other Type of Waiver Months	OTH_WVR_1115_MOS
Other Waiver Type Months	OTH_WVR_TYPE_MOS

4. Money Follows the Person (MFP)

The MFP supplemental file contains one record for each beneficiary (identified by each unique combination of MSIS_ID and STATE_CD) who participated in an MFP program in any month during the calendar year. Variables include monthly Money Follows Person (MFP) Participant indicators (variables called MFP_IND_01-12). The file also contains constructed variables that identify the last known MFP status for the beneficiary within the year (Table 8). Additional information regarding the MFP Program is available on the Medicaid.gov website[10].

Table 8. DE Money Follows Person (MFP) Supplemental File – Annual Summary Variables

Description	SAS variable name
Money Follows Person (MFP) - Lives with Family or Non-Participant Code	MFP_LVS_WTH_FMLY_CD
Money Follows Person (MFP) - Qualified Residence Code	MFP_QLFYD_RSDNC_CD
Money Follows Person (MFP) - Qualified Institution Code	MFP_QLFYD_INSTN_CD
Money Follows Person (MFP) - Reinstitutionalized Reason Code	MFP_RINSTLZD_RSN_CD
Money Follows Person (MFP) - Participation Ended Reason Code	MFP_PRTCPTN_END_RSN_CD

5. Health Home and State Plan Options (SPO)

The Health Home and SPO supplemental file contains one record for each beneficiary (identified by each unique combination of MSIS_ID and STATE_CD) who participated in a Health Home or who was enrolled in a SPO, or had a Health Home Chronic condition in any month during the calendar year. Variables include the name and provider identifier for the Health Home, as well as monthly Health Home Program Participation Indicator variables (called HLTH_HOME_PGM_IND_01 – 12). The file also has monthly variables that indicate enrollment in state plan options for each month, including: Community First Choice, 1915i, 1915j, 1932a, and 1937 Alternative Benefit Plan. Additional information regarding these plans is available on the Medicaid.gov website[11]. The eight Health Home Chronic Condition indicators are: asthma, diabetes, heart disease, mental health, substance abuse, HIV/AIDs, overweight and other conditions.

[10] MFP is explained on the Medicaid.gov website: https://www.medicaid.gov/medicaid/ltss/money-follows-the-person/index.html. Accessed 10/07/2019.

[11] For example, Community First Choice (CFC) is described at: https://www.medicaid.gov/medicaid/hcbs/authorities/1915-k/index.html . Accessed 09/25/2019.

6. Disability and Need

The Disability and Need supplemental file contains one record for each beneficiary (identified by each unique combination of MSIS_ID and STATE_CD) who met one of the following criteria:

- met the requirements for Long-Term Services and Supports (LTSS) program eligibility;
- was eligible through the conception to birth option;
- participated in other government assistance programs; or
- had any of the following – a Home and Community-Based Services (HCBS) chronic condition, a lock-in provider, a disabling condition, or third-party liability coverage.

Within the file, there are ten HCBS Chronic Condition indicators: aged, physical disability, intellectual disability, developmental disability, other disability, autism, mental illness, brain injury, HIV/AIDs, and technology dependent medically fragile[12]. If beneficiaries are "locked-in" to specific providers to better monitor services and reduce unnecessary or inappropriate utilization, there are variables for identification of up to three lock-in providers (provider ID -LCKIN_PRVDR_ID_1- 3, each with an associated provider type code - LCKIN_PRVDR_TYPE_CD_1 – 3).

There are also monthly disability indicators for particular activities, as well as monthly variables related to enrollment in the following programs: Birth to Conception, Social Security Disability Insurance (SSDI), Supplemental Security Income (SSI), Temporary Assistance for Needy Families (TANF), and Third Party Liability (TPL) Insurance Coverage. Refer to Table 9.

Table 9. DE Disability and Need Supplemental File – Monthly Variables

Monthly Disability Variables - Description	**SAS variable name**[13]
Disability Indicator - Difficulty Concentrating (one variable for each month)	DSBLTY_DFCLTY_CNCNTRTNG_IND_01 - 12
Disability Indicator - Difficulty Walking (one variable for each month)	DSBLTY_DFCLTY_WLKG_IND_01 - 12
Disability Indicator - Difficulty Dressing or Bathing (one variable for each month)	DSBLTY_DFCLTY_DRSNG_BTHNG_IND_01 - 12
Disability Indicator - Difficulty Running Errands Alone (one variable for each month)	DSBLTY_DFCLTY_ERNDS_IND_01 -12

Monthly Program Eligibility Variables - Description	**SAS variable name**
Birth to Conception Indicator (one variable for each month)	BIRTH_CNCPTN_IND_01 -12
Social Security Disability Insurance (SSDI) Indicator (one variable for each month)	SSDI_IND_01 -12
Supplemental Security Income (SSI) Indicator (one variable for each month)	SSI_IND_01 -12
Supplemental Security Income (SSI) Status Code (one variable for each month)	SSI_STUS_CD_01 -12
Supplemental Security Income (SSI) State Supplement Code (one variable for each month)	SSI_STATE_SPLMT_CD_01 -12

[12] Additional information is in DQ Brief #7041, "Identifying and Benchmarking the Number of Medicaid 1915(c) Waiver Participants in 2016" on the ResDAC website.
[13] These are monthly variables (01 thru 12).

Monthly Program Eligibility Variables - Description	SAS variable name
Temporary Assistance for Needy Families (TANF) Cash Code (one variable for each month)	TANF_CASH_CD_01 -12
Third Party Liability (TPL) Insurance Coverage Indicator (one variable for each month)	TPL_INSRNC_CVRG_IND_01 -12
Third Party Liability - Other Coverage Indicator (one variable for each month)	TPL_OTHR_CVRG_IND_01-12

Finally, there are monthly indicators for the overall level of care required to meet an individual's needs and to determine LTSS program eligibility (CARE_LVL_STUS_CD_01-12). Related variables include up to three LTSS providers (LTSS_PRVDR_ID_1 – 3) during the year, and the associated monthly values for the Level of Care Codes for each of these providers (LTSS_LVL_CD_1_01 – 12 for provider 1 through LTSS_LVL_CD_3_01 - 12 for provider 3). In addition to the monthly level of care codes, there is a variable with the latest value for the level of care codes 1, 2 and 3 during the year (e.g., the Long-Term Services & Supports Provider 2 Level of Care Code - Latest in Year is called LTSS_LVL_CD_2_LTST).

Chapter 3. TAF Research File Claims

States submit fee-for-service (FFS) claims, managed care encounter claims, and supplemental payments for Medicaid, Medicaid-expansion CHIP, and Separate CHIP to T-MSIS. As stated in Chapter 1, Section A, Source Data, the TAF claims files include all "final action" Medicaid and CHIP service records for a given year (i.e., all T-MSIS claims CMS determined to be final, as of the TAF creation date). The claims included in these files are active, final-action, non-voided, and non-denied claims[14] (except for IL, see Chapter 4, Section C.3, Known Issue with Illinois Claims.

It is important to note that there is significant state-by-state variability in the quality of claims and encounter data; refer to the Data Quality briefs on the ResDAC website and information later in Chapter 4 of this user guide regarding TAF Data Quality (see, for example, Chapter 4, Section C, Considerations for Selecting States for Analysis.

Only claims with a service end date in the month and year, along with their associated claim line records, are included in the TAFs for that year.

The TAF claims files are available for four settings:

1. Inpatient (IP),
2. Long term care (LT),
3. Other services (OT), and
4. Pharmacy (RX).

The RIF data for each of the settings is partitioned into header and trailer records, which we refer to as the "header" and "line" files. In addition, there are "occurrence" code files for IP, LT and OT. The header file contains the core/basic service record, while the line file may contain multiple line items for a corresponding header record[15]. The header and line file records for a service are linked using the CLM_ID, which is a variable added by the CCW specifically to link claim header and line records.

The lines associated with a header claim are numbered sequentially using the sequential claim line number (LINE_NUM). Similarly, the occurrence code records associated with a header claim are numbered sequentially using the occurrence code sequence (variable called OCRNC_CD_SEQ). The LINE_NUM and OCRNC_CD_SEQ variables are both added by the CCW to facilitate ease of use of the files. Later in this Chapter we present some examples of how to associate lines or occurrence code records with the header claim.

A. Structure of Claims Files

1. Header Files

The header file contains claim header information such as the BENE_ID, MSIS_ID, claim type code, service begin (or admission) date, service end (or discharge) date, adjudication date, provider ID, managed care

[14] "Non-denied" claims mean they were not denied at the header level; there may be denied lines in the line file – i.e. the claim was not completely denied, however some lines for these claims may be denied.
[15] It is possible, although not common, to have a header claim with no associated line records (i.e., nothing to join to within the line file).

plan ID, waiver ID and type code, diagnosis codes, as well as the diagnosis-related group (DRG), and up to six procedure codes with associated dates (for IP records).

The number of diagnosis code fields varies by file:

- for IP there are up to 12 DX codes,
- for LT there are up to five,
- for OT there are up to two diagnosis codes, and
- RX has none.

Claim-level payment information is also present for all FFS claims; however the Center for Medicaid and CHIP Services (CMCS) at CMS has required redaction of some payment fields for managed care claims due to the proprietary and confidential nature of this information (refer to Appendix C – Managed Care Redacted Payment Fields). Redacted payment fields on the RIF include: Medicaid paid amount, other insurance paid amounts (Medicare or third party), and beneficiary liability amounts (deductibles and coinsurance). When summing expenditures using the claim files (either header or line) it is important to keep in mind that the redacted managed care expenditures will not be reflected in your totals.

2. Line Files

Records in the line files are associated with the header claim (i.e., a particular CLM_ID). There will never be "orphan" lines, (i.e. line records without an associated header claim record), although it is possible for there to be a header claim record without any associated line records.

Line files have a separate record for each revenue center or ambulatory procedure (note that for IP claims the header record accommodates major procedures designated with a procedure code). Although the procedure is often documented using an ICD-9 or ICD-10-PCS code, states may use their own codes (refer to variables that identify the procedure code system/ nomenclature used - PRCDR_CD_SYS_1 - 6). Revenue centers represent institutional cost centers for which separate charges are billed. For example, inpatient hospital claims may include revenue centers for emergency department (ED), intensive care, physical therapy, laboratory, pharmacy, blood, imaging, etc. For non-institutional settings, claim lines may represent distinct procedures (e.g., CPT or HCPCS codes), or national drug codes (NDCs).

It is common to use the line-level detail if you are interested in the OT or RX claims because the line record contains important information to identify the services that were delivered. The type of service code (TOS_CD) is assigned at the line level to describe the service (as opposed to the MAX claims, where the type of service code, MSIS_TOS, is assigned at the header claim level).

The line files include the beginning and ending dates of service for the line, as well as information about the service provider and payment. Payment fields on the RIF line are also redacted for managed care claims (refer to Appendix C – Managed Care Redacted Payment Fields).

An example of using the lines associated with the header claim can be seen in Table 10.

Table 10. Example of Header and Line Records for an OT Claim

BENE_ID	header or line file	CLM_ID	LINE_NUM	SRVC_BGN_DT (on header)	SRVC_END_DT (on header)	Procedure Code or Revenue Center Code	Type of service code	BLG_PRVDR_ID
000001	header	123456	n/a	14MAY2016	14MAY2016	n/a	n/a	ABC117
000001	line	123456	1	(LINE_SRVC_BGN_DT) 14MAY2016	(LINE_SRVC_END_DT) 14MAY2016	99213	003	line (SRVC_PRVDR_ID) 007
000001	line	123456	2	14MAY2016	14MAY2016	36415	003	007
000001	line	123456	3	14MAY2016	14MAY2016	G0467	003	007
000885	header	987654	n/a	11JUL2016	11JUL2016	n/a	n/a	QRS999999
000885	line	987654	1	(LINE_SRVC_BGN_DT) 11JUL2016	(LINE_SRVC_BGN_DT) 11JUL2016	line revenue center code = 0260	002	line (SRVC_PRVDR_ID) 2224444
000885	line	987654	2	11JUL2016	11JUL2016	0335	002	2224444

In this example there is a header claim followed by three associated line records for BENE_ID 000001. The line records indicate this beneficiary visited the Federally Qualified Health Center (FQHC, where type of service code=003), where he is an established patient (procedure code, CPT4=99213) receiving a complete physical, with collection of lab specimen (CPT code=36415). For the second example (Table 10), the line records indicate BENE_ID 000885 is receiving chemotherapy at a hospital outpatient department. Rather than procedure codes, this provider used revenue center codes to indicate the services provided (revenue center code 0260=IV infusion pump and 0335=Radiology, therapeutic chemotherapy IV).

Revenue codes and HCPCS codes are well-established standardized medical billing nomenclatures, while the Medicaid type of service code nomenclature is created by CMS. Revenue center and type of service valid values can be found in the TAF Research Files codebook at http://www.ccwdata.org/web/guest/data-dictionaries.

3. Occurrence Code Files

Occurrence codes provide information when there is a special condition which may affect how the state paid for the service (e.g., date of an injury). Occurrence codes are submitted for about 46% of 2016 IP records, 6% of 2016 LT records and 2% of 2016 OT records. The RX file does not have an Occurrence Code file. A claim may have up to 10 occurrence codes and dates, although only a subset of claims will have a record in the occurrence code files. To use this file, link to the header claim file using the CLM_ID. The occurrence records associated with a header claim are numbered sequentially using the occurrence code sequence (OCRNC_CD_SEQ).

Occurrence codes may identify a significant event/date relating to an IP or LT stay (e.g., accident, medical condition, insurance-related, service-related, etc.). The corresponding start date for the occurrence is listed (variable called OCRNC_CD_START_DT), along with the ending date, when applicable (OCRNC_CD_END_DT). It is common for the occurrence start date to precede the dates of service on the claim. An example of occurrence codes is presented in Table 11.

Table 11. Example of Occurrence Code Records

BENE_ID	header or occurrence file	CLM_ID	OCRNC_CD_SEQ	OCRNC_CD	SRVC_BGN_DT (on header)	SRVC_END_DT (on header)
000421	header	123456	n/a		27JUL2016	02AUG2016
000421	occurrence	123456	1	A2	(OCRNC_CD_START_DT) 01FEB2016	
000737	header	789012	n/a		10MAY2016	13MAY2016
000737	occurrence	789012	1	11	(OCRNC_CD_START_DT) 09MAY2016	(OCRNC_CD_END_DT) 09MAY2016
000737	occurrence	789012	2	18	01JAN2014	

In this example there is an occurrence code (OCRNC_CD) associated with the header claim for BENE_ID 000421 (A2), which means that the occurrence date represents the effective date for the insurance policy. The occurrence code start date is February 1, 2016 and the claim admission date is July 27, 2016.

BENE_ID 000737 also had occurrence codes associated with the header claim (Table 11). The first occurrence code (11) refers to the date of the onset of symptoms/illness. The second occurrence code (i.e., where OCRNC_CD_SEQ=2) is 18, which refers to the date of retirement for the patient/beneficiary. Occurrence codes are a standardized medical billing nomenclature; valid values can be found in the TAF Research Files codebook at http://www.ccwdata.org/web/guest/data-dictionaries.

B. Content and Description of Claims RIFs

Claims for various types of providers and services are located in different files. This section describes the contents of each of the four claims files. A complete list of variables is available on the ccwdata.org website in the TAF Research Files Record Layout, and detailed variable descriptions are provided in the associated Codebook[16].

[16] Please see https://www.ccwdata.org/web/guest/data-dictionaries for the TAF Research Files Codebooks.

1. Inpatient (IP)

This file is for inpatient hospital services. Unlike the MAX files, where claims were summarized into "stays", the TAF is at the claim-level, and a single hospital stay may be represented by more one claim (e.g., for lengthy hospitalizations). The lines for these claims will typically have the type of service code (TOS_CD) equal to '001'. Other type of service codes typically expected in other settings may also be found in inpatient lines. Refer to Table 12 and Appendix D – Type of Service Codes by File; the TAF Codebook contains an exhaustive list of TOS_CD values. For additional information regarding the TOS_CDs in the data files, refer to DQ Brief #5141, "Variation in the Use of Type of Service Codes in 2016," on the ResDAC website.

Key information on these claims includes diagnosis codes (up to 12 are allowed) and the corresponding DRG. For each diagnosis code, there is an indicator to flag whether the diagnosis was present on admission (POA). Up to six (6) major procedures may be coded, with the associated procedure date. The procedure code may be a standard ICD-9 or ICD-10 code, or it may be a state-specific procedure code.

Table 12. Examples of Type of Service Code by File Type

TOS Code*	TOS Description	IP	LT	OT	RX
001	Inpatient hospital services, other than services in an institution for mental diseases	X			
002	Outpatient hospital services			X	
003	Rural health clinic services			X	
006	Technical laboratory services			X	
008	Technical radiological services			X	
009	Nursing facility services for individuals age 21 or older (other than services in an institution for mental disease)		X		
010	Early and periodic screening and diagnosis and treatment (EPSDT) services			X	
012	Physicians' services			X	
014	Outpatient substance abuse treatment services			X	
016	Home health services - Nursing services			X	
028	Clinic services			X	
029	Dental services			X	
030	Physical therapy services (when not provided under home health services)			X	
033	Prescribed drugs				X
034	Over-the-counter medications				X
036	Medical equipment/prosthetic devices			X	X

* A full list of Type of service codes is available in the Codebook; you may also refer to Appendix D – Type of Service Codes by File. A TOS code value may appear in more than one file type.

2. Long Term Care (LT)

This file is for long-term care institutional claims, including nursing facilities, intermediate care facility services for individuals with intellectual disabilities, mental health facility services, and independent (free-standing) psychiatric wings of acute care hospitals.

The lines for these claims will often have the type of service code (TOS_CD) equal to '009', and '044'-'048', although other type of service codes may also be present. Refer to Appendix D – Type of Service Codes by File. Note that similar types of services may be provided in a non-institutional, or home-based setting, and these claims will typically be found in the OT file.

Key information on the LT claims includes diagnosis codes (up to five are allowed), provider information, program and waiver information, and payment information (redacted for managed care encounter records).

3. Other Services (OT)

This file contains information for a broad range of other services; basically all services that are not classified strictly as IP, LT or RX. OT records can represent bills from either facilities or providers. These two broad groups use slightly different formats for their bills so that the information varies a little between claims based on a facility bill compared to claims based on a provider bill. Services in the OT TAF include but are not limited to: physician services, outpatient hospital services, dental services, other physician services (e.g., chiropractors, podiatrists, psychologists, optometrists, etc.), clinic services, laboratory services, X-ray services, sterilizations, home health services, personal support services, and managed care capitation payments.

Key information on the OT claims includes diagnosis codes (up to two are allowed), provider, and payment information (redacted for managed care encounter records). Much of the information regarding the service will be found on the claim lines – including procedures, drugs, etc. Refer to example in Table 10. Note that this file may also include many drugs, such as chemotherapy and other infused drugs or vaccines.

The lines for these claims will have a broad range of type of service codes (TOS_CD). Refer to Table 12 and Appendix D – Type of Service Codes by File. A provider may submit a claim with multiple types of services, which means that each line of the claim may have a different type of service code. Revenue center codes (REV_CNTR_CD) may appear on the lines for facility claims, whereas procedure codes (LINE_PRCDR_CD) are often found on the lines for providers.

4. Pharmacy (RX)

This file contains information for both prescription and over-the-counter drugs that are covered by Medicaid or CHIP. In general, only the claims for drugs or other products (e.g., bandages or diabetic test strips) provided by a pharmacy will be found in this file.

The lines for these claims will typically have the TOS_CD equal to '033' or '034', although other type of service codes may also be present. Refer to Table 12 and Appendix D – Type of Service Codes by File.

For RX claims, information regarding the drug dispensed is on the line file (i.e., the national drug code, or NDC). Some drugs will appear in the OT file, such as infused drugs (including chemotherapies) and vaccines. In addition, drugs provided as part of a hospitalization are not included in this file since they are not paid separately and are instead bundled as part of the hospital stay. Similarly there are rarely drug claims in the LT file since most of those services are provided within an institutional facility.

Note that for some states, there may be a header RX claim without any associated line records. Additional information and tips for working with this file appear in Chapter 4, Section A, Linkage Variables and Joining Files.

C. Key Concepts and Variables in Claims RIFs

The TAFs are very large. It is possible that, for a given study, smaller analytical files may be desirable in order to easily manipulate the data for the intended results. When creating an analytic file, it may be prudent to subset the claims files to retain only the key claims and variables of interest. For example, you may wish to only include claims for the study population that meet your criteria (e.g., particular states, duration of enrollment, type of program/benefit, age group or who received a particular type of service [TOS_CD]).

Some key TAF variables are highlighted below. Additional information regarding all the variables in the file can be found in the TAF Research Files codebook on the ccwdata.org website.

1. Type of claim

Claim type code – This header claim field (CLM_TYPE_CD) is used to identify what kind of payment is covered on the claims record. This field distinguishes between claims that are for Medicaid or Medicaid-expansion, S-CHIP, and other types of claims. Values are listed in Table 13 below.

Table 13. Claim Type Codes Associated with Medicaid, Medicaid Expansion and S-CHIP

Description	Medicaid or Expansion Claims	S-CHIP Claims	Other Claims
Fee-for-service (FFS) claim	1	A	U
Capitated payment	2	B	V
Encounter record	3	C	W
Supplemental payment	5	E	Y

Note that service tracking records (i.e., where CLM_TYPE_CD=4, D, X) and claims with a missing CLM_TYPE_CD are not included in the TAF Research Files since they do not represent discrete services provided to patients and have no corresponding diagnosis or procedure information.

2. Reason for Service

The claims records include fields to help identify the diagnoses, procedures and/or medications dispensed in an ambulatory setting. Various claims coding systems are used to define these services. Refer to Table 15.

- **Diagnosis** - The number of diagnosis code fields varies by file: 1) for IP there are up to 12 diagnosis codes, 2) for LT there are up to five, 3) for OT there are up to two diagnosis codes, and 4) RX has none.
    - On October 1, 2015, CMS, in accordance with the Health Insurance Portability and Accountability Act (HIPAA), converted from the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) to the 10th Revision, Clinical Modification (ICD-10-CM) and Procedure Coding System (ICD-10-PCS).
    - Regardless of when a claim was submitted for payment, services that occurred prior to October 1, 2015 were billed using ICD-9 codes and all services provided on or after

October 1, 2015 utilized ICD-10 codes[17]. States may have required providers to submit claims or encounter records with ICD-10 codes on a slightly different schedule. Within the TAF RIFs there are header file variables (DGNS_VRSN_CD_1 - X) that indicate whether a diagnosis code is ICD-9, ICD-10 or something else).

In September 2015, the ICD-9 diagnosis code was used almost 100% of the time in the IP header file, for example. In October 2015, only 1.96% of IP header claims used ICD-9 codes and 98.03% used ICD-10 codes (refer to Table 14).

Table 14. ICD Diagnosis Version Code in IP Header File

DGNS_VRSN_CD_1	September 2015	October 2015	January 2016
1 = ICD-9	99.89%	1.96%	1.59%
2 = ICD-10	0.11%	98.03%	98.40%
3 = Other	0.00 %	0.01 %	0.01 %

- **Procedures** – when examining procedures codes, you will also need to use the procedure code system variables (one for each of the procedure codes) called PRCDR_CD_SYS_1-X (six in IP header file; also one on the OT line – LINE_PRCDR_CD_SYS). See Table 15.
- **DRG** – states are not required by CMS to use the Medicare DRG classification system. When a DRG code (variable called DRG_CD) appears on the claim, there is a corresponding DRG code system variable (called DRG_CD_SYS and associated description – DRG_DESC). See also Table 15.

Table 15. Claims Coding Systems by File type

Type of Code	IP	LT	OT	RX
ICD-9/ICD-10-CM diagnosis code*	X	X	X	
Diagnosis-related group (DRG)**	X			
Revenue center code	X	X	X	
Procedure code†	X	X	X	
National Drug Code (NDC)		X	X	X

* ICD-9 or ICD-10 codes may be used, depending on the service date.

**DRG may be a state code (not Medicare DRG). Refer to the variable on the claim with the DRG called DRG_CD_SYS.

† Procedure codes may be submitted using a variety of formats including ICD-9/ICD-10-PCS procedure, Current Procedural Terminology (CPT) code, Healthcare Common Procedure Coding System (HCPCS) code, or may be a state code; refer to the PRCDR_CD_SYS_1-X variable.

3. Dates of Service

Service begin and end dates – The SRVC_BGN_DT is the beginning date of service on the claim/encounter record covering services rendered to the beneficiary (i.e., the claim record covers services starting on this date). The service begin date may not always correspond with the first date the beneficiary received services (i.e., it is not always equal to admission date), if services were provided over a long period of time, such as for lengthy LT stays.

[17] CMS. ICD-10 webpage. https://www.cms.gov/medicare/coding/icd10/index.html. Accessed 09/25/2019.

Similarly, the claim ending date of service (SRVC_END_DT) is the last day on the claim billing statement for services rendered to the beneficiary (i.e., the claim record covers services through this date). Note that this date may not always correspond with the last date the beneficiary received services (i.e., it is not always equal to discharge date). RX claims do not have a service begin or end date on the header claim file; the prescription fill date indicates the date of service (variable called RX_FILL_DT).

For LT and OT claims, the header claim variable SRVC_END_DT is utilized for claim selection. In the case of IP and RX claims, header claim variables discharge date (DSCHRG_DT) and prescription fill date (RX_FILL_DT) are utilized, respectively.

For some states the header claim SRVC_END_DT or DSCHRG_DT is not always populated. When this happens in the IP, LT and OT RIFs, the CCW utilizes a hierarchy of service date fields to populate the SRVC_END_DT (this same hierarchy is used for TAF selection also). For example, if the discharge date in an IP claim is not populated, CCW evaluates the maximum line service end date for the IP claim to populate the header SRVC_END_DT variable. The OT, LT, and IP claim types each have their own hierarchy. The variable SRVC_END_DT_CD identifies the service date variable value used to impute the SRVC_END_DT in the claim header file.

For IP, LT and OT files there are also service begin/end dates on the line file (variables called LINE_SRVC_BGN_DT and LINE_SRVC_END_DT). These may be of value if you are interested in the beginning/ending date of service for the service line representing a home health visit, for example, and typically are used for many kinds of OT services analysis. RX line file claims do not have a line service begin or end date; use the prescription fill date variable (variable called RX_FILL_DT, which CCW copies from the RX header and includes in the RX line file).

Admission and discharge dates – IP and LT claims have admission (ADMSN_DT) and discharge (DSCHRG_DT) dates. There are times when a service date on the claim was missing. As mentioned above, in this circumstance, CCW uses the TAF selection algorithm mentioned previously to impute a SRVC_END_DT value so that we may ensure claims are associated with the appropriate date for the data file. We include a variable (called the service end date code - SRVC_END_DT_CD) to identify when and how the date was imputed; additional details are in the codebook.

Procedure date – the IP header claim has dates for each of the six (6) procedure code fields. The date identified in the PRCDR_CD_DT_# applies to the procedure in the corresponding PRCDR_CD_#. In addition, the OT line file has a procedure date variable (LINE_PRCDR_CD_DT) that corresponds to the line procedure code (LINE_PRCDR_CD).

4. Provider Information

The TAF claims have fields to accommodate several different types of providers. In addition, a variety of different types of identification numbers may be utilized by states. States may use a Medicaid-specific provider identification number (variables with the *_PRVDR_ID suffix), and/or CMS provider identifiers (variables with the *_PRVDR_NPI suffix). Refer to specific variable examples in Table 16. There are fields to accommodate provider identifiers on both the header and line files. Additionally, there are taxonomy codes, which are a set of codes used to identify the type of provider and areas of specialization for health care providers.

The table below depicts the various types of provider IDs in the Claims TAFs (Table 16).

Table 16. Provider ID Variables in Claims

Provider	File	State-assigned Identification Number	CMS National Provider Identifier (NPI)	Taxonomy Code	Provider Type	Specialty Code
Admitting	IP, LT	ADMTG_PRVDR_ID	ADMTG_PRVDR_NPI	ADMTG_PRVDR_TXNMY_CD	ADMTG_PRVDR_TYPE_CD	ADMTG_PRVDR_SPCLTY_CD
Billing	All	BLG_PRVDR_ID	BLG_PRVDR_NPI	BLG_PRVDR_TXNMY_CD	BLG_PRVDR_TYPE_CD*	BLG_PRVDR_SPCLTY_CD
Referring	IP, LT, OT	RFRG_PRVDR_ID	RFRG_PRVDR_NPI	RFRG_PRVDR_TXNMY_CD**	RFRG_PRVDR_TYPE_CD	RFRG_PRVDR_SPCLTY_CD
Directing	OT	n/a	DRCTNG_PRVDR_NPI	DRCTNG_PRVDR_TXNMY_CD	n/a	n/a
Supervising	OT	n/a	SPRVSNG_PRVDR_NPI	SPRVSNG_PRVDR_TXNMY_CD	n/a	n/a
Prescribing	RX	PRSCRBNG_PRVDR_ID	PRSCRBNG_PRVDR_NPI	n/a	n/a	n/a
Dispensing	RX	DSPNSNG_PRVDR_ID	DSPNSNG_PRVDR_NPI	n/a	n/a	n/a
Line – Servicing	IP,LT and OT Line	SRVC_PRVDR_ID	SRVC_PRVDR_NPI	SRVC_PRVDR_TXNMY_CD	SRVC_PRVDR_TYPE_CD	SRVC_PRVDR_SPCLTY_CD
Line – Operating	IP Line	n/a	OPRTG_PRVDR_NPI	n/a	n/a	n/a

* The RX file type does not have the billing provider type code on the header record.
** OT is the only file type that has the referring taxonomy code.

Some health home services can be found in the OT files. For these records, the OT header file includes information related to the health home provider NPI (HLTH_HOME_PRVDR_NPI) and the health home entity name (HLTH_HOME_ENT_NAME).

5. Billing and Payment Information

CCW is required to employ claim-level redaction of managed care payment fields when the payer is a managed care plan (refer to Appendix C – Managed Care Redacted Payment Fields). Managed care claims are identified using the field CLM_TYPE_CD and the values '3', 'C', or 'W'. All other values are considered non-managed care and do not require redaction. The other RIF fields are populated for these claim records (i.e., no additional suppression is applied for the managed care records).

The claims RIFs have several types of variables related to payment:

- The amount Medicaid paid for the service
- The amount the beneficiary had to pay for the service (in the form of deductibles or copayments)
- The amount paid by Medicare or a third party for the service.

States may also make additional payments to providers or managed care plans, such as disproportionate share hospital (DSH) payments or other lump sum payments. These claims do not represent claims for services provided to individual beneficiaries; therefore these claims are NOT included in the RIF.

6. Program/Plan/Waiver Types

The header claims (for all four claim types) have fields to capture other key applicable information including:

- Program Type Code (PGM_TYPE_CD) – such as Early and Periodic Screening and Diagnosis and Treatment (EPSDT) or Money Follows the Person (MFP)
- Managed Care Plan Identification Number (MC_PLAN_ID)
- Waiver Type Code (WVR_TYPE_CD), which may indicate a variety of 1915 and 1115 waivers or state plan options, and the associated Waiver Identification Number (WVR_ID)
- Indicator Insured is Covered by Another Plan (Not Medicare or Medicaid) (OTHR_INSRNC_IND)
- 1115(A) Demonstration Participation Indicator (SECT_1115A_DEMO_IND) for CMMI demonstrations

States submit their Medicaid expenditure reports to CMS using a standard CMS-64 form[18], and they submit CHIP expenditures on the CMS-21 form. The claim line files for each of the four claim types include fields to categorize the service in terms of the CMS-64 form and CMS-21 form:

- CMS-64 Form Category of Service for the Paid Claim (XIX_SRVC_CTGRY_CD)
- CMS-21 Form Category of Service for the Paid Claim (XXI_SRVC_CTGRY_CD)
- CMS-64 Form Code for Federal Reimbursement (CMS_64_FED_CTGRY_CD)

[18] Additional information regarding these reports can be found on the Medicaid.gov website (for example - https://www.medicaid.gov/medicaid/finance/state-expenditure-reporting/index.html). Accessed 09/25/2019.

Chapter 4. Using TAF Research Files

Researchers are encouraged to become familiar with the various CMS Data Quality briefs and Medicaid and CHIP resources[19] prior to working with the data files. For those who have worked with MAX files previously, there are major differences in the TAFs that should be considered prior to undertaking analyses. This section of the document is intended to provide some helpful hints and considerations for working with the TAF Research Files.

A. Linkage Variables and Joining Files

As discussed earlier in the document, the same BENE_ID may be associated with more than one MSIS_ID and STATE_CD if the person was enrolled in Medicaid or CHIP in more than one state. In fact, the BENE_ID may also be associated with more than one MSIS_ID in the same state if the eligible has multiple non-contiguous spells of Medicaid or CHIP enrollment. As such, when joining the DE enrollment files together (e.g. the DE Base to any of the DE supplemental files), use the combination of MSIS_ID and STATE_CD to ensure that you are associating the correct supplemental records to the selected base record. A visual depiction of the rationale for using MSIS_ID and STATE_CD for this linkage is in Table 17.

Table 17. Examples of Joining DE Base file with Supplemental File(s)

BENE_ID (DE Base)	MSIS_ID (DE Base)*	STATE_CD (DE Base)	BENE_ID (Suppl. File)	MSIS_ID (Suppl. File)	STATE_CD (Suppl. File)	WVR_1915C_MOS (Suppl. File)**
000123	000ABCDE1357	NH	000123	000ABCDE1357	NH	9
000123	0762468WXYZ	NY	000123	0762468WXYZ	NY	1
000123	R99134ST8877	MA	000123	R99134ST8877	MA	0

* Within the data files, MSIS_ID is 32 alphanumeric characters; shortened in this example for brevity.
** This example assumes we are using the Waiver Supplemental File.

In this example, the same person (BENE_ID 000123) is enrolled in Medicaid or CHIP in three states during the year. To ensure you are joining the base enrollment record for that person to the associated DE supplemental file records (e.g., wavier program enrollment), you must use the MSIS_ID and STATE_CD. In this example, since there is more than one DE base record for the BENE_ID, it is not possible to accurately join these files using the BENE_ID – you risk joining the supplemental DE records to the wrong base records.

If you are using multiple years of data and wish to identify the same person longitudinally, we recommend using the BENE_ID, since the MSIS_ID for a beneficiary may change over time. These linkage variables (BENE_ID, MSIS_ID and STATE_CD) appear in all TAF research files.

To link the header claim with associated lines, use the unique claim ID (variable called CLM_ID), which is present on all header and line files. The line file detail records associated with the claim are numbered sequentially (variable called line number LINE_NUM). A visual depiction of the linkage between the header claim and associated lines is shown in Table 10. You can also join the header claim to the occurrence code

[19] ResDAC https://www.resdac.org/.

file (which is only available/populated for a subset of IP, LT and OT claims) with the CLM_ID. Remember that the associated occurrence code records are numbered sequentially (OCRNC_CD_SEQ).

Unlike the other three claim types, it is common to see RX header claims with no associated lines because the header claim alone is sufficient to describe the prescription/dispensing of a drug. There may also be RX claims with more than one associated line; this generally means the medication dispensed was a compound drug (i.e., more than one active ingredient/NDC).

The data files can be very large; therefore we recommend creating an analytic file with a subset of variables to make files more manageable. We recommend that the users do NOT include all variables on the header and line files when combining the data.

B. Summarizing Data

We recommend using the BENE_ID, when available, to identify all claims for the beneficiary for the year. If there is not a BENE_ID for the record, then use the combination of MSIS_ID and STATE_CD. Similarly, when summarizing enrollment, although you join the DE Base and supplemental files using the MSIS_ID and STATE_CD, we recommend summarizing the data using the BENE_ID. Some scenarios are presented below, and a visual depiction is shown in Table 18:

- *Summarizing services across states*. If looking at all of the services received by a beneficiary during the year regardless of the state (in cases where beneficiaries were in more than one state), then use the BENE_ID. If you are interested in services for a beneficiary in a single state, then use BENE_ID within a STATE_CD.
- *De-duplicating enrollment within a state.* If you want to identify the distinct number of beneficiaries in a state, use the BENE_ID (since someone might have more than one MSIS_ID within a state).
- *Enrollment across states*. If you are interested in the number of months of Medicaid enrollment across years for a beneficiary, use the BENE_ID.

Table 18. Examples of Summarizing Data for Beneficiaries

BENE_ID	MSIS_ID*	STATE_CD	Months of Medicaid	# Inpatient claims
000123	000ABCDE1357	NH	3	1
000123	0762468WXYZ	NY	4	1
000123	R99134ST8877	MA	5	2
Total			**12**	**4**

* Within the data files, MSIS_ID is 32 alphanumeric characters; shortened in this example for brevity.

The example in Table 18 shows that BENE_ID 000123 was enrolled in Medicaid in three different states during the year. Accordingly, there are three different MSIS_ID and STATE_CD combinations. Looking only within a single state, an inaccurate ascertainment of Medicaid enrollment and services would be obtained for the beneficiary. The BENE_ID allows you to look across states for a more complete view of enrollment and services.

C. Considerations for Selecting States for Analysis

States have flexibility with program implementation and may expand the scope of benefits and how services are delivered (e.g., through use of waivers, etc.). Thus, it is important to understand how a specific state's program operates (e.g., managed care or FFS) as well as the range of benefits for the state(s) of interest. A variety of sources may be helpful in understanding state Medicaid and CHIP programs, including the Medicaid.gov website[20]. Additionally, data quality varies by state, topic, and year. The data quality resources featured on the ResDAC website can provide insight into these issues.

States also have flexibility regarding when to convert from MSIS to T-MSIS, although all states converted to T-MSIS by October 2015.

1. State cutover to T-MSIS

The CCW did not create TAF research files for all states for 2014 or 2015 (refer to Table 1). A state needed a full calendar year of T-MSIS data to be included in the TAFs. The states with TAFs, by year, is located in Appendix B – State Cutover to T-MSIS and Availability of MAX and TAF Research Files. If there is a not a TAF for the state(s) of interest, a MAX file should be available in most cases.

2. State CHIP and Eligibility Groups

The claims TAFs are not limited to services provided through the Medicaid and CHIP programs. States may have claims for services provided to other (or expansion) populations, depending on state-specific enrollment criteria. See Medicaid.gov for "State Overviews" https://www.medicaid.gov/state-overviews/index.htmlState Overviews | Medicaid.gov

Within the DE TAF you can identify populations enrolled in Medicaid, CHIP, expansion populations or other state programs. Refer to Data Tip 1. Furthermore, within the claims RIFs, you can identify claims that were for Medicaid, CHIP, or other services using the claim type code (variable called CLM_TYPE_CD).

3. Known Issue with Illinois Claims

States have varying levels of data quality (refer to the ResDAC website for state data quality information). For those wishing to use claims from Illinois, some special processing is needed. Refer to Data Tip 2 below.

Data Tip 2. Illinois – Incremental Claims

Unlike the claims files from other states, the claims files submitted by Illinois are not limited to final action records. When adjudicating claims, IL Medicaid processes incremental claims - rather than "replacement" claims. This means that adjustments to claims (e.g., credits and debits) must be considered in conjunction with the original claim. For IL, all transactional claims/encounter records are included in the RIF. Additional information and guidance is available on the ResDAC website in the document "TAF Technical Guidance: How to Use Illinois Claims Data."

[20] A resource for national and state Medicaid information is Medicaid.gov. See, for example: https://www.medicaid.gov/medicaid/program-information/index.html . Accessed 09/25/2019.

D. Cautions in Working with TAF RIFs

1. Claims without Enrollment

There are some Medicaid and CHIP claims for which there is no corresponding enrollment record for the beneficiary receiving the service. As a result, CMS has created a "dummy" enrollment record for these beneficiaries, however, there is no Medicaid/CHIP enrollment information in the dummy record. The "dummy" records will only include a MSIS-ID and state code. No additional information (birth date, gender or enrollment dates) will be found on the record. The dummy enrollment records are identified by a missing eligibility data indicator (MISG_ELGBLTY_DATA_IND) value of 1.

As discussed in Chapter 3, Section C.1, Type of claim, some service tracking records are not included in RIF (e.g., DSH payments and bundled capitated payments). In some instances, a corresponding "dummy" enrollment record may appear in the DE base file for these records. If you are counting beneficiaries/enrollees in a state, we recommend removing the dummy enrollment records from the analysis. Do not simply count MSIS_IDs for each state because this will result in an over count of beneficiaries in the state.

2. Date Variable Transformations

T-MSIS source data submitted by states may occasionally include date values that are implausibly early. Due to CCW software defaults, the T-MSIS date values prior to 1/1/1600 are recoded to 12/31/1599, with the exception of 1/1/1000, which is set to null in the RIFs.

E. TAF Data Quality Briefs and Other Supplemental Materials

The TAF Data Quality briefs are designed to support TAF RIF data users. A variety of topics and variable-specific questions are analyzed by state to underscore the usability of the data. This information is intended to be used to design analyses, and to inform analytic work. There are dozens of Data Quality briefs, organized into "themes" (e.g., enrollment, managed care, provider information, etc.). In addition to the issue briefs, there are state and topic "snapshots" designed to provide a quick, high-level evaluation of the quality of the state's data by topic area. A link to all data quality materials can be found on the ResDAC website (https://www.resdac.org/).

Chapter 5. Receiving CCW Data

The TAFs are provided to academic researchers and certain government agencies, to conduct approved research studies under a Data Use Agreement (DUA). The CCW Medicaid and CHIP data contain identifiable information, and are subject to the Privacy Act and other Federal government rules and regulations (see ResDAC web site for information on requesting Medicare and CHIP data http://www.resdac.org/). Anyone seeking approval for access to the Medicaid and CHIP data must ensure that their research proposal can be reliably supported by the data.

External researchers have two options for accessing the data files – they may access them directly from the CCW within the Virtual Research Data Center (VRDC), or they can have their data shipped. For CCW/VRDC users, their access is initiated by the CCW team upon receipt of the approved DUA and payment.

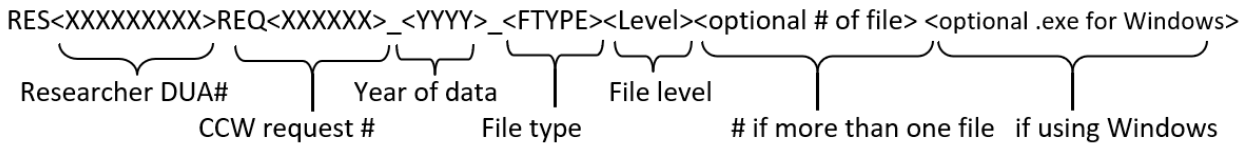
For researchers who request that their data be shipped, once the DUA is in place and payment for the files has been received, data files are prepared and shipped to the requestor on either a USB external hard drive, DVD, or flash drive. These data files are packaged as encrypted self-decrypting archive (SDA) files. The decryption password is sent to you electronically via email. When you receive the data package (via hard drive or DVD/CD), copy them from the shipping media to your local workspace. Note some data shipped on a hard drive can be decrypted on that hard drive, depending on the size of the data files. Using the password provided to you via email, follow the Decryption Instructions enclosed in the data package. Each SDA contains the data file(s), SAS® code and a file transfer summary (.fts) file which can be used to verify the data was read in correctly.

A. Format

The files that are delivered to the researcher are organized in the following format. There will be several folders, each of which contains multiple files. The folders are:

XXXX (folder with your CCW data request #)
Extract File Documentation

There may be additional folders if you have requested additional types of data. All of your data files are located within the folder with your CCW data request # (see Table 20). There is a separate sub-folder for each year of data you requested. The naming convention for data files is as follows:



For example, if your DUA # was 77777, your CCW request number was 012345 and you obtained 2016 T-MSIS OT data, your folders and data files would look like this:

012345
2016
READ_ME_FIRST_REQ012345_2016.txt
res77777req012345_2016_TAFOTH.exe

The types of claims (or other data files) are identified as follows. See Table 19.

Table 19. T-MSIS RIF Enrollment and Claim File Names

Enrollment File Type	Base/Supplemental Files	File Level
DE Base File	demog_elig_base	Base
Eligibility Dates	demog_elig_dates	Supplement
Disability and Need	demog_elig_disability	Supplement
Health Home and State Plan Options (SPO)	demog_elig_hh_spo	Supplement
Managed Care Enrollment	demog_elig_mngd_care	Supplement
Money Follows Person (MFP)	demog_elig_mny_flw_prsn	Supplement
Waiver Program Enrollment	demog_elig_waiver	Supplement
Claims File Type	**Header/Line/Occurrence Files**	**File Level**
Inpatient (IP)	inpatient_header inpatient_line inpatient_occurrence	Header Line Occurrence
Long Term Care (LT)	long_term_header long_term_line long_term_occurrence	Header Line Occurrence
Other Services (OT)	other_services_header other_services_line other_services_occurrence	Header Line Occurrence
Pharmacy (RX)	rx_header rx_line	Header Line

If the files are extremely large they may be divided into two or more files, in which case there would be a sequential number at the end of the file name - such as “001”, “002”, to enumerate how many files of this type you receive, for example:

res77777req012345_2016_taf_other_services_header_001.dat
res77777req012345_2016 taf_other_services_line_001.dat
res77777req012345_2016 taf_other_services_line_002.dat

Table 20. CCW Output Package - Data File Folder

File	File Description
READ_ME_FIRST_REQXXX_YYYY.txt	This is a text file that describes the files contained in the output package. File Name Example: READ_ME_FIRST_REQ012345_2016.txt

Claims Files

File Name	File Description	Unit of Analysis
res<DUA number>_req<XXX>_<YYYY>_ taf_other_services_header_<# files>	Other Services (OT) Header Claim File	Claim
res<DUA number>_req<XXX>_<YYYY>_ other_services_line_<# files>	Other Services Line File	Line Detail
res<DUA number>_req<XXX>_<YYYY>_ OT _OCCR	Other Services Occurrence Code File	Occurrence Detail*

* The Occurrence Code File is delivered for each of the claims file types (IP, LT and OT; it is not applicable to the RX file).

B. Content

Within each of the yearly data folders is a README file, which you will want to read first. It is a text file that describes the files contained in the output package.

All of the data files are contained within executable files (Self-Decrypting Archive [SDA]). **You will need to enter a password to extract each file.** Additional details regarding the data encryption and extraction process are in section C, below.

After you extract the data files, you should compare your record count to the control counts that the CCW obtained in producing your data file. These control counts are in the *.fts file. There is a separate .fts for each data file. The data files are in fixed column flat files. You can use whatever analytic software you choose. For convenience, we have included SAS® read-in files. In addition to the raw data files, the following files are generated by each executable in the output package (Table 21).

Table 21. Files Contained within SDAs

File Name	File Description
<file name>.fts	For each extracted data file there will be a corresponding transfer summary file. The names of these files will correspond with the data file name [e.g., res<DUA number>_req<XXX>_<YYYY>_ other_services_line. fts]. This file transfer summary files contain: • File name • File source • File transfer mode • Row length • File transfer format • # Columns • # Rows • File size
<file name>_<level>_v6.sas	Program to read data into a SAS version 6.x environment. For example, the file inpatient_header_claims_read_v6.sas reads the inpatient header claims data into a SAS version 6.x environment.
<file name>_<level>_v8.sas	Program to read data into a SAS version 8.x environment.
<file name>_<level>.dat	Naming convention for *.dat files is in Table 19

Extract File Documentation

Code Reference Sets.xls – describes the ICD-9 or 10 diagnosis and procedure codes, HCPC codes, revenue center and other codes in the data files.

Decryption Instructions.pdf – instructions for decrypting/uncompressing the data files.

Tips on Getting Started with Data

Chapter 6. Further Assistance with Medicaid and CHIP Data

There are multiple Help Desk resources available to assist with technical and content questions and problems. Please see below to determine which Help Desk to contact.

A. Ordering Data Files, General Contents Questions, and DUA

The Research Data Assistance Center (ResDAC) is the CMS contractor that assists researchers in obtaining the CCW data. The ResDAC website provides links to descriptions of the available CMS data, request procedures, supporting documentation, and other helpful resources. In particular, the ResDAC site hosts a range of TAF Data Quality briefs and other supplemental materials that have been produced to describe state-level quality variation in commonly used variables and concepts within the TAF data files. Visit the ResDAC website at (http://www.resdac.org) for additional information.

Researchers can reach ResDAC by:

www.resdac.org
Email: resdac@umn.edu
Phone: 1-888-973-7322

If additional CMS data (data not available from the CCW) is required to meet research objectives, or the researcher has any questions about other data sources, the researcher can review all available CMS data by visiting the ResDAC website and contacting ResDAC for further assistance.

B. Questions After Data are Received or Access is Granted from the CCW

If you have already received the TAF RIFs and have questions or concerns regarding the data files, access to the VRDC or need assistance with the media, please contact the CCW:

www.ccwdata.org
Email: ccwhelp@gdit.com
Phone: 1-866-766-1915

All RIF record layouts and codebooks, as well as this User Guide, are posted on the ccwdata.org website.

Appendix A - List of Acronyms

CCW – Chronic Conditions Data Warehouse

CHIP – Children's Health Insurance Program both M-CHIP and S-CHIP (unless specified otherwise)

CMCS – Center for Medicaid and CHIP Services

CMS – Centers for Medicare & Medicaid Services

CPT – Current Procedural Terminology

DE – Demographic and Eligibility File

DQ – Data Quality (briefs)

DRG – Diagnosis Related Group

DUA – Data Use Agreement

FPL – Federal Poverty Level

FFS – Fee-For-Service claims

HCBS – Home and Community-Based Services

HCPCS – Healthcare Common Procedure Coding System

ICD-9 (or 10) – International Classification of Diseases, Ninth Revision (or Tenth)

IP – Inpatient File

LT – Long Term Care File

LTSS – Long-Term Services and Supports

MAX – Medicaid Analytic eXtract data files

MAX-T – MAX file produced with T-MSIS data

M-CHIP – Medicaid Expansion Children's Health Insurance Program

MFP – Money Follows the Person

MSIS – Medicaid Statistical Information System

NDC – National Drug Code

NPI – National Provider Identifier; CMS-required identification number

OT – Other Services File

ResDAC – Research Data Assistance Center

RIF(s) – Research Identifiable File(s)

RX – Pharmacy File

S-CHIP – Separate Children's Health Insurance Program

SDA – Self-Decrypting Archive

SPO – State Plan Options

SSI – Supplemental Security Income

TAF(s) – T-MSIS Analytic File(s)

T-MSIS – Transformed Medicaid Statistical Information System

TPL – Third Party Liability

VRDC – Virtual Research Data Center

Appendix B – State Cutover to T-MSIS and Availability of MAX and TAF Research Files

State	ANSI State Code (STATE_CD)	First TMSIS Submission	Medicaid Data Source for 2014	Medicaid Data Source for 2015	Medicaid Data Source for 2016
Alabama	AL	Jan-14	TAF	TAF	TAF
Alaska	AK	Oct-13	TAF	TAF	TAF
Arizona	AZ	Oct-14	MAX-T	TAF	TAF
Arkansas	AR	Apr-15	MAX-T	MAX-T	N/A
California	CA	Oct-15	MAX	MAX-T	TAF
Colorado	CO	Oct-11	TAF	TAF	TAF
Connecticut	CT	Apr-15	MAX-T	MAX-T	TAF
Delaware	DE	Jan-14	TAF	TAF	TAF
District of Columbia	DC	Jan-14	TAF	TAF	TAF
Florida	FL	Oct-13	TAF	TAF	TAF
Georgia	GA	Oct-15	MAX	MAX-T	TAF
Hawaii	HI	Oct-14	MAX-T	TAF	TAF
Idaho	ID	Oct-15	MAX	MAX-T	TAF
Illinois	IL	Jan-14	TAF	TAF	TAF
Indiana	IN	Oct-14	MAX-T	TAF	TAF
Iowa	IA	Oct-15	MAX	MAX-T	TAF
Iowa CHIP	IA	Jan-14	N/A	N/A	TAF
Kansas	KS	Jan-13	TAF	TAF	TAF
Kentucky	KY	Jul-14	MAX-T	TAF	TAF
Louisiana	LA	Oct-15	MAX	MAX-T	TAF
Maine	ME	Jan-14	TAF	TAF	TAF
Maryland	MD	Jan-14	TAF	TAF	TAF
Massachusetts	MA	Oct-14	MAX-T	TAF	TAF
Michigan	MI	Oct-15	MAX	MAX-T	TAF
Minnesota	MN	Oct-15	MAX	MAX-T	TAF
Mississippi	MS	Oct-15	MAX	MAX-T	TAF
Missouri	MO	Oct-15	MAX	MAX-T	TAF
Montana	MT	Jan-14	TAF	TAF	TAF
Montana TPA	MT	Jan-16	N/A	N/A	N/A
Nebraska	NE	Jan-14	TAF	TAF	TAF
Nevada	NV	Jan-14	TAF	TAF	TAF
New Hampshire	NH	Jan-14	TAF	TAF	TAF
New Jersey	NJ	Oct-15	MAX	MAX-T	TAF
New Mexico	NM	Jan-14	TAF	TAF	TAF
New York	NY	Jul-15	MAX-T	MAX-T	TAF
North Carolina	NC	Jul-13	TAF	TAF	TAF

State	ANSI State Code (STATE_CD)	First TMSIS Submission	Medicaid Data Source for 2014	Medicaid Data Source for 2015	Medicaid Data Source for 2016
North Dakota	ND	Jan-14	TAF	TAF	TAF
Ohio	OH	Oct-14	MAX-T	TAF	TAF
Oklahoma	OK	Oct-14	MAX-T	TAF	TAF
Oregon	OR	Jul-15	MAX-T	MAX-T	TAF
Pennsylvania	PA	Oct-15	MAX	MAX-T	TAF
Pennsylvania CHIP	PA	Jan-14	N/A	N/A	TAF
Puerto Rico	PR	Jan-15	N/A	N/A	N/A
Rhode Island	RI	Oct-12	TAF	TAF	TAF
South Carolina	SC	Jul-14	MAX-T	TAF	TAF
South Dakota	SD	Oct-15	MAX	MAX-T	TAF
Tennessee	TN	Oct-15	MAX	MAX-T	TAF
Texas	TX	Jul-14	MAX-T	TAF	TAF
Utah	UT	Oct-15	MAX	MAX-T	TAF
Vermont	VT	Oct-15	MAX	MAX-T	TAF
Virginia	VA	Apr-14	MAX-T	TAF	TAF
Virgin Islands	VI	Jan-17	N/A	N/A	N/A
Washington	WA	Jan-15	MAX-T	TAF	TAF
West Virginia	WV	Oct-15	MAX	MAX-T	TAF
Wisconsin	WI	Jan-14	TAF	TAF	TAF
Wyoming	WY	Oct-15	MAX	MAX-T	TAF
Wyoming CHIP	WY	Oct-15	N/A	N/A	N/A

Notes:
MAX = MAX produced using MSIS
MAX-T = MAX produced using TMSIS, at least partly
TAF = T-MSIS Analytic Files
N/A = No data available

Appendix C – Managed Care Redacted Payment Fields

For the TAF RIF, Managed Care payment fields are restricted. Below are the 36 fields in the claim types that will have restricted access, according to CMCS.

- IP header
  1. Total Medicaid paid amount (MDCD_PD_AMT)
  2. Total Medicare deductible amount (MDCR_DDCTBL_PD_AMT)
  3. Total Medicare coinsurance amount (MDCR_COINSRNC_PD_AMT)
  4. Total billed amount (BILLED_AMT)
  5. Total allowed amount (MDCD_ALOWD_AMT)

- IP line
  6. Medicaid paid amount (LINE_MDCD_PD_AMT)
  7. Charge amount (REV_CNTR_CHRG_AMT)
  8. Allowed amount (LINE_MDCD_ALOWD_AMT)

- LT header
  9. Total Medicaid paid amount (MDCD_PD_AMT)
  10. Total Medicare deductible amount (MDCR_DDCTBL_PD_AMT)
  11. Total Medicare coinsurance amount (MDCR_COINSRNC_PD_AMT)
  12. Total billed amount (BILLED_AMT)
  13. Total allowed amount (MDCD_ALOWD_AMT)
  14. Aggregate room and board (MDCD_ACMDTN_PD_AMT)
  15. Aggregate ancillary amount (MDCD_ANCLRY_PD_AMT)

- LT line
  16. Medicaid paid amount (LINE_MDCD_PD_AMT)
  17. Charge amount (REV_CNTR_CHRG_AMT)
  18. Allowed amount (LINE_MDCD_ALOWD_AMT)

- RX header
  19. Total Medicaid paid amount (MDCD_PD_AMT)
  20. Total Medicare deductible amount (MDCR_DDCTBL_PD_AMT)
  21. Total Medicare coinsurance amount (MDCR_COINSRNC_PD_AMT)
  22. Total billed amount (BILLED_AMT)
  23. Total allowed amount (MDCD_ALOWD_AMT)

- RX line
  24. Medicaid paid amount (LINE_MDCD_PD_AMT)
  25. Medicare deductible amount (LINE_MDCR_DDCTBL_PD_AMT)
  26. Medicare coinsurance amount (LINE_MDCR_COINSRNC_PD_AMT)
  27. Billed amount (LINE_BILLED_AMT)
  28. Allowed amount (LINE_MDCD_ALOWD_AMT)

- OT header
  29. Total Medicaid paid amount (MDCD_PD_AMT)
  30. Total Medicare deductible amount (MDCR_DDCTBL_PD_AMT)
  31. Total Medicare coinsurance amount (MDCR_COINSRNC_PD_AMT)

32. Total billed amount (BILLED_AMT)
33. Total allowed amount (MDCD_ALOWD_AMT)

- OT line
  34. Medicaid paid amount (LINE_MDCD_PD_AMT)
  35. Billed amount (LINE_BILLED_AMT)
  36. Allowed amount (LINE_MDCD_ALOWD_AMT)

These 36 fields will only have values suppressed for Managed care claims. Managed care type claims are identified using the field CLM_TYPE_CD = '3', 'C', 'W'. All other values are considered non-managed care and do not require restriction. All other RIF fields are populated for these claim records (i.e., no additional suppression is applied for the managed care records).

Appendix D – Type of Service Codes by File

Type of Service Code (SAS name in RIF is TOS_CD) by File Type. A TOS code value may appear in more than one file type. For additional information regarding the TOS_CDs in the data files, refer to DQ brief #5141, "Variation in the Use of Type of Service Codes"on the ResDAC website.

TOS Code	TOS Description	IP	LT	OT	RX
001	Inpatient hospital services, other than services in an institution for mental diseases	X			
002	Outpatient hospital services			X	
003	Rural health clinic services			X	
004	Other ambulatory services furnished by a rural health clinic			X	
005	Professional laboratory services, Technical laboratory services			X	
006	Technical laboratory services			X	
007	Professional radiological services			X	
008	Technical radiological services			X	
009	Nursing facility services for individuals age 21 or older (other than services in an institution for mental disease)		X		
010	Early and periodic screening and diagnosis and treatment (EPSDT) services			X	
011	Family planning services and supplies for individuals of child-bearing age			X	X
012	Physicians' services			X	
013	Medical and surgical services of a dentist			X	
014	Outpatient substance abuse treatment services			X	
015	Medical or other remedial care or services, other than physicians' services, provided by licensed practitioners within the scope of practice as defined under State law			X	
016	Home health services - Nursing services			X	
017	Home health services - Home health aide services			X	
018	Home health services - Medical supplies, equipment, and appliances suitable for use in the home			X	X
019	Home health services - Physical therapy provided by a home health agency or by a facility licensed by the State to provide medical rehabilitation services			X	
020	Home health services - Occupational therapy provided by a home health agency or by a facility licensed by the State to provide medical rehabilitation services			X	
021	Home health services - Speech pathology and audiology services provided by a home health agency or by a facility licensed by the State to provide medical rehabilitation services			X	
022	Private duty nursing services			X	
023	Advanced practice nurse services			X	
024	Pediatric nurse			X	
025	Nurse-midwife service			X	
026	Nurse practitioner services			X	
027	Respiratory care for ventilator-dependent individuals			X	
028	Clinic services			X	
029	Dental services			X	
030	Physical therapy services (when not provided under home health services)			X	
031	Occupational therapy services (when not provided under home health services)			X	
032	Speech, hearing, and language disorders services (when not provided under home health services)			X	
033	Prescribed drugs				X

TOS Code	TOS Description	IP	LT	OT	RX
034	Over-the-counter medications				X
035	Dentures			X	
036	Medical equipment/prosthetic devices			X	X
037	Eyeglasses			X	
038	Hearing Aids			X	
039	Diagnostic services			X	
040	Screening services			X	
041	Preventive services			X	
042	Well-baby and well-child care services as defined by the State.			X	
043	Rehabilitative services			X	
044	Inpatient hospital services for individuals age 65 or older in institutions for mental diseases		X		
045	Nursing facility services for individuals age 65 or older in institutions for mental diseases		X		
046	Intermediate care facility (ICF)/ Intermediate Care Facilities for individuals with Intellectual Disabilities (IIDICF)/ Individuals with Intellectual Disabilities (IID) services		X		
047	Nursing facility services, other than in institutions for mental diseases		X		
048	Inpatient psychiatric services for individuals under age 21		X		
049	Outpatient mental health services, other than Outpatient substance abuse treatment services. This TOS includes services furnished in a State-operated mental hospital and including community-based services.			X	
050	Inpatient substance abuse treatment services and residential substance abuse treatment services.		X	X	
051	Personal care services			X	
052	Primary care case management services			X	
053	Targeted case management services			X	
054	Case Management services other than those that meet the definition of primary care case management services or targeted case management services			X	
055	Care coordination services			X	
056	Transportation services			X	
057	Enabling services			X	
058	Services furnished in a religious nonmedical health care institution	X			
059	Skilled nursing facility services for individuals under age 21		X		
060	Emergency hospital services	X		X	
061	Critical access hospital services – OT			X	
062	HCBS - Case management services			X	
063	HCBS - Homemaker services			X	
064	HCBS - Home health aide services			X	
065	HCBS - Personal care services			X	
066	HCBS - Adult day health services			X	
067	HCBS - Habilitation services			X	
068	HCBS - Respite care services			X	
069	HCBS - Day treatment or other partial hospitalization services, psychosocial rehabilitation services and clinic services (whether or not furnished in a facility) for individuals with chronic mental illness			X	

TOS Code	TOS Description	IP	LT	OT	RX
070	HCBS - Day Care			X	
071	HCBS - Training for family members			X	
072	HCBS - Minor modification to the home			X	
073	HCBS - Other services requested by the agency and approved by CMS as cost effective and necessary to avoid institutionalization			X	
074	HCBS - Expanded habilitation services - Prevocational services			X	
075	HCBS - Expanded habilitation services - Educational services			X	
076	HCBS - Expanded habilitation services - Supported employment services, which facilitate paid employment			X	
077	HCBS-65-plus - Case management services			X	
078	HCBS-65-plus - Homemaker services			X	
079	HCBS-65-plus - Home health aide services			X	
080	HCBS-65-plus - Personal care services			X	
081	HCBS-65-plus - Adult day health services			X	
082	HCBS-65-plus - Respite care services			X	
083	HCBS-65-plus - Other medical and social services			X	
084	Sterilizations	X		X	
085	Prenatal care and pre-pregnancy family planning services and supplies			X	X
086	Other Pregnancy-related Procedures	X		X	
087	Hospice services			X	
088	Any other health care services or items specified by the Secretary and not excluded under regulations			X	
089	Disposable medical supplies			X	X
090	Critical access hospital services – IP	X			
091	Skilled care – hospital residing	X			
092	Exceptional care – hospital residing	X			
093	Non-acute care – hospital residing	X			
115	Residential care			X	
119	Capitated payments to HMOs, HIOs, or PACE plans			X	
120	Capitated payments for primary care case management (PCCM)			X	
121	Premium payments for private health insurance			X	
122	Capitated payments to prepaid health plans (PHPs)			X	
123	Disproportionate share hospital (DSH) payments	X		X	
127	Indian Health Service (IHS) - Family Plan			X	X
131	Drug Rebates			X	X
132	Supplemental payment – inpatient	X			
133	Supplemental payment – nursing		X		
134	Supplemental payment – outpatient			X	
135	Electronic health record (EHR) payments to provider	X		X	