Chronic Condition Data Warehouse

Your source for national CMS Medicare and Medicaid research data

Medicaid Analytic eXtract Files (MAX)

User Guide

June 2019

Version 2.3

Overview

Medicaid is a state-administered benefit with guidance/requirements and shared funding from the federal government and individual state governments (Title XIX of 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, whether pregnant, 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 disabled children living at home. For more information on the Medicaid program, see the Centers for Medicare & Medicaid Services (CMS) website (www.cms.gov).

Each state compiles information regarding enrollment, service utilization, and payment in their Medicaid Management Information System (MMIS). CMS provides the state with a data dictionary to map their MMIS data elements into the Medicaid Statistical Information System (MSIS). Each state provides final, reconciled MSIS data files to CMS in the MSIS format. CMS and its contractors compile the various state MSIS data files into a uniform data structure for each calendar year and make these annual state segment files available to researchers as the Medicaid Analytic eXtract (MAX) data files. The MAX data are created to support research and policy analysis.

Researchers will find useful information about file content by state and year by reviewing the MAX Data Validation Tables and Data Anomalies Reports under the Medicaid Analytic eXtract (MAX) General Information on Data section of the CMS website (https://www.cms.gov/MedicaidDataSourcesGenInfo/07_MAXGeneralInformation.asp).
Another useful resource for researchers is the MAX Chartbook, which provides a data appendix summarizing enrollee demographic, eligibility characteristics, Medicaid service use, and Medicaid expenditures at national and state levels. MAX Chartbooks are available on the CMS website as well. The three versions currently available include data from 2002, 2004, or 2008. Many of these documents, in addition to statistical resources, can also be referenced on the ResDAC web site at (http://www.resdac.org/Medicaid/data_available.asp).

The MAX data are contained in the CMS Chronic Conditions Data Warehouse (CCW) and are provided to academic researchers and certain government agencies, which have been approved under a Data Use Agreement (DUA) to obtain MAX data for research purposes. The CCW MAX 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 MAX data).

The CCW MAX files are linked by a unique CCW person identifier (the beneficiary identifier; variable called BENE_ID), allowing researchers to analyze information across the continuum of care, within and across years, without using actual identifying information (e.g., MSIS_ID, Social Security number, or Medicare Health Insurance Claim number for dual eligibles). This unique identifier follows an enrollee across states so that if the same person was enrolled in Medicaid in

more than one state, those records may be combined using the BENE_ID. Record layouts for the CCW MAX files are available on the ccwdata.org website (https://www.ccwdata.org/web/guest/data-dictionaries).

There are six types of data files available for CCW MAX data, including the following:
1) Person Summary File (PS)
2) Hospital Inpatient File (IP)
3) Long Term Care File (LT)
4) Prescription Drug File (RX)
5) Other Services (OT)
6) Medicaid Enrollee Supplemental File (MESF), which is designed to augment the MAX PS data

This guide provides users with information that may be helpful in understanding and requesting the CCW MAX data.

Change Log

Date	Changed By	Changes	Version
6/01/2019	K. Schneider	Change log created. Including Sickle Cell Disease (SCD)	2.3

Table of Contents

Chapter 1. General Information about MAX Files

The original MAX data are extracted from MSIS data for each state and compiled by CMS and its contractors. The MAX development process combines MSIS initial claims, interim claims, voids, and adjustments for a given service into a final action event. Unlike fiscal-based MSIS quarterly files, MAX data are organized into annual calendar year files, based on date of service. Since it is necessary to allow for the delay between service delivery dates and claims adjudication dates, the availability of MAX data for a particular time period lags behind that of the MSIS data. States may finalize their MSIS files at different points in time, which can lead to variation in the timing/production of the MAX files.

Historically, the MAX PS File was developed prior to the CCW; therefore it lacked some of the key features of CCW data files that are designed to aid in the study of quality of care and health of the population. This was remedied when CCW began disseminating an additional file designed to supplement the information in the MAX PS. The Medicaid Enrollee Supplemental File (MESF) consists of two optional components called "segments". These became available in 2014. The two file segments are the:
1) Conditions segment and
2) NDI segment.

These file segments are described in Chapter 2. Investigators may elect to purchase these segments along with their MAX PS data.

States have some flexibility with regard to benefit design, and may request waivers from CMS in order to offer programs in a particular manner. The specific types of Medicaid services offered and requirements for eligibility may vary somewhat by state and by year as well. For some researchers, the nuances of the benefit design may be important factors to consider when examining Medicaid across different states (for more information on state waiver and plan amendment information, see (http://medicaidbenefits.kff.org/index.jsp or http://www.cms.gov/MedicaidGenInfo/01_Overview.asp). For additional information on the creation of the MAX data files, see the CMS website (http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/index.html?redirect=/MedicaidDataSourcesGenInfo/07_MAXGeneralInformation.asp).

Since the MAX data contain individually identifiable data, they are protected under the Privacy Act. They are available for approved research activities only through a DUA with CMS.

For each calendar year beginning with 1999, the six MAX data file types available are:

1. **Person Summary (PS)**

Person-level information regarding Medicaid-eligibles who have enrolled in a state Medicaid program or a Medicaid-expansion Child Health Insurance Program (CHIP) in the year, whether or not they used any services

2. **Medicaid Enrollee Supplemental File (MESF) –** CCW developed two types of files which can be used to augment information in the MAX PS.
   - ➢ **Conditions segment** - The CCW pre-codes specific chronic condition categories for duals and the Medicaid only population using fee-for-service claims data to demonstrate a pattern of treatment as evidence of the presence of a condition.
   - ➢ **NDI segment** - The cause of death, according to official death certificate data purchased from the National Death Index (NDI) through the Centers for Disease Control and Prevention (CDC). This segment contains only the subset of enrollees who had a date of death in the MAX PS file; or if the person was dually enrolled in Medicare, a date of death in Medicare or Social Security Administration records. This file is available through 2013.

3. **Inpatient Hospital (IP)**

Contains complete stay records for enrollees who used inpatient services. Records include diagnoses, procedures, discharge status, length of stay, and payment amounts.

4. **Long Term Care (LT)**

Includes service data from four types of long-term care facilities that serve Medicaid enrollees, as well as data for similar services provided in community-based settings.

5. **Prescription Drug (RX)**

Contains final action records of prescription and over-the-counter drugs paid by Medicaid. Records include a National Drug Code (NDC) to indicate what drug product was used, service date and payment information.

6. **Other Services (OT)**

Includes outpatient, physician and professional services; hospice, home health, lab/X-ray, durable medical equipment (DME), premium payments etc., and all Medicaid records not reported in any other MAX service files.

These are calendar year files which are produced each year as the Medicaid source data are available. There is a lag time for CMS to produce these data files – while all of the states allow providers to submit and reconcile claims for the service year, then while each Medicaid agency finalizes and reconciles their Medicaid MSIS data with CMS; data are generally available for most states about 3 years after the service year. MAX files are currently available for 1999 through 2009. Note the record layouts for each file type changed slightly in 2005 (i.e., 1999-2004 file layouts differ from 2005 forward), and new values were introduced in 2006. The MAX PS record layout changed for 2007. Historical MAX record layouts are available on the CMS website (http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-

Systems/MedicaidDataSourcesGenInfo/MAXGeneralInformation.html). All current record layouts for the CCW MAX files are on the ccwdata.org website. Notable changes will be described within each file type below.

The PS File contains a record for each Medicaid enrollee (i.e., one record per MSIS_ID per state).

- The **MESF Conditions segment** contains information for each enrollee in the MAX PS for the year; however, the MESF uses the unique beneficiary identifier (BENE_ID) in combination with the state (variable called STATE_CD) as the level of aggregation for each row. The determination regarding the presence or absence of treatment for conditions is made at the unduplicated person level, which means if the person was enrolled in Medicaid in more than one state during the year there will be more than one row of data in the Conditions segment, however the rows of data will be identical. Additional details regarding level of aggregation for the Conditions segment appear in the MAX PS Chapter; additional information regarding the BENE_ID appears in the "Assignment of a Beneficiary Identifier" section below, and also in the Chapter describing the contents of the MAX PS.

- The **MESF NDI segment** contains one row for each BENE_ID with a date of death in either the MAX PS file, or, if the person was dually enrolled in Medicare – we included those with a Medicare-reported date of death (available through 2013).

For service claim types LT and OT, each record of data represents a claim for services; for IP, each record represents an inpatient stay; for RX, each record represents a claim for a prescription or over- the-counter drug product. Additional information regarding data dictionaries, data element lists, validation reports, and previously documented data anomalies reports are available on the CMS website at (https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAXGeneralInformation.html).

CCW published a technical guidance document to assist MAX data file users in getting started with analyzing the data files. There are a variety of helpful hints for working with the data, as well as SAS® analytic code examples for some common types of analyses. See "Getting Started with MAX Data Files – A Technical Guidance Paper" on the CCW website (https://www.ccwdata.org/web/guest/technical-guidance-documentation).

Assignment of a Beneficiary Identifier
In the CCW, all Medicare enrollment, Medicare Part A claims, Part B claims, Part D events, Minimum Data Set (MDS)/Outcome and Assessment Information Set (OASIS) and Medicaid Analytic Extract (MAX) data records are assigned a beneficiary identifier (BENE_ID). The BENE_ID identifies a unique beneficiary (common person) across the distinct data sources, thereby allowing researchers to track all services for a beneficiary over time. Medicaid enrollees who are also enrolled in the Medicare program (or "dual eligibles") are assigned the same BENE_ID for both MAX and Medicare records within the CCW. This allows for ease of analysis across programs, data types (including CCW assessment data), and years. The unique CCW BENE_ID is specific to the CCW and is not applicable to any other identification system or data source.

To construct the CCW BENE_ID, the CMS CCW team developed an internal cross-reference file consisting of historical Medicaid and Medicare enrollment information using CMS data sources such as the Enterprise Cross Reference (ECR) file. When a new MAX PS file is received, the MSIS_ID, STATE_CD, SSN, DOB, Gender and other beneficiary identifying information is compared against the historical enrollment file. If there is a single record in the historical enrollment file that "best matches" the information in the MAX PS record, then the BENE_ID on that historical record is assigned to the MAX PS 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 MAX PS record. For any given year, approximately 7% to 8% of MAX PS records have a BENE_ID that is null. Once a BENE_ID is assigned to a MAX PS record for a particular year (with the exception of those assigned to a null value), it will not change. When a new MAX PS file is received, CCW attempts to reassign those with missing BENE_IDs.

The BENE_ID from the MAX PS file is assigned to each respective claim record (IP, LT, RX, OT), as well as each record in the MESF. This is the main linkage key to use when joining data files. Note for PS records with a null BENE_ID, the linkage should be by MSIS_ID and STATE_CD.

Chapter 2. Medicaid Person Summary (PS) Files and Supplemental Files

The PS file contains one record for every individual enrolled during the calendar year represented by the file. Note that it is possible that a BENE_ID may appear more than once – for example, if the person received a second MSIS_ID either in the same state or a different state. The PS file contains enrollment, including waiver enrollment, demographic, and summary utilization data. Note that this file includes people enrolled in Medicaid, CHIP, and Medicaid enrollees covered by Section 1115 Demonstrations or other Medicaid waiver programs.

Currently, the MAX data files contain cost and use information for M-CHIP (Medicaid expansion CHIP) since Federal matching funds are used. However, S-CHIP (optional state CHIP) enrollment reporting is not required, therefore, there is no S-CHIP utilization information in the MAX files.

The specific types of Medicaid services offered and requirements for eligibility may vary somewhat by state and year. For some researchers, the nuances of the benefit design may be important factors to consider when examining Medicaid enrollment, utilization, and payment summaries in the PS file. This between-state variation may also be important to consider when making comparisons between states.

The data dictionary for the MAX Files is located on the CCW website (https://www.ccwdata.org/web/guest/data-dictionaries).

The PS File contains a rich array of information about the demographic, eligibility, and summary utilization and payment information. Important types of information include:

- *Demographic Information* – date of birth, sex, race/ethnicity, date of death, state, county and ZIP code

- *Eligibility Group Information* – the monthly MAX uniform eligibility codes, monthly enrollment status, and whether there was dual eligibility (i.e., concurrent Medicare enrollment). Also includes information regarding waivers, premium payments, third party payers, and managed care enrollment

- *Inpatient Hospital (IP) Utilization Summary* – total covered days and stays

- *Institutional Long Term Care (LTC) Utilization Summary* – annual aggregated covered days by setting: mental health, inpatient psychiatric, ICF/MR, nursing facility, and total LTC days

- *Claims Payment Summary* – information from the other four record types (IP, LT, RX and OT) for each of six program types (Family Planning [FP], Federally Qualified Health Centers [FQHC], Home and Community-Based Care for Aged and Elderly Disabled [HCBCA], Home and Community-Based Services [HCBCS], Indian Health Service [IHS], and Rural Health Clinic

[RHC]). Also includes information regarding total records, total Medicaid payments, encounter records etc., for the 31 types of service

The primary payer (e.g., Medicaid versus Medicare or another third party payer) may vary depending on the service; however, if the person is also covered by Medicare, it is the primary payer for all Medicare-covered services. Medicare enrollees may qualify for "full" or only "partial" Medicaid coverage, which will affect the payment amounts observed on the MAX claims (note: see the monthly variables called *eligible restricted benefits flag* for more information about these categories).

The basis of eligibility (BOE) and maintenance assistance status (MAS) fields are reported data elements in MSIS, the original source for MAX data. The code values for these two data elements are concatenated into a single data element in MAX, known as the MAX Uniform Eligibility Code variable. MAS is in the first position, and the BOE is in the second position of this data element.

Changes beginning with the 2005 MAX PS file and values added to the 2006 file include the addition of monthly crossover codes to allow for more precise determination of the Medicaid/Medicare dual coverage (i.e., "crossover" coverage).

Changes beginning with the 2007 MAX PS file included the addition of the Social Security Administration (SSA) death master file date of death (*effective April 2014, this variable is no longer available for release per CMS*), removal of the quarterly crossover (dual) codes; and renumbering of CMS MAX data elements between 26 and 35.

Medicaid Enrollee Supplemental File

Investigators may purchase either of the following MESF data file segments to augment information in the MAX PS. The two CCW-created enrollee segments contain person-level information; however, there is not a 1:1 match between the MAX PS file and either of the MESF segments. For the NDI segment, records are at the BENE_ID-level when this information is available, rather than the MSIS_ID by STATE_CD-level (i.e., the NDI segment does not contain duplicate rows of information in cases where a unique person was represented in the MAX PS by more than one enrollment record). For the Conditions segment, there is a row of data for each state where the BENE_ID is present for the year; and if there is not a BENE_ID for the enrollee, there is a record for each MSIS_ID within each state. Table 1 details the record counts for each of the files. The data dictionaries for the MESF Segments are available on the CCW website (https://www.ccwdata.org/web/guest/data-dictionaries).

Table 1. MAX PS and MESF segment File sizes.

Year	MAX PS	Conditions	NDI*
1999	43,587,106	43,574,193	673,013
2000	46,334,479	46,320,468	691,653
2001	50,078,314	50,061,046	713,993
2002	55,063,855	55,032,144	737,159
2003	57,638,888	57,605,187	763,439

Year	MAX PS	Conditions	NDI*
2004	60,244,145	60,214,750	755,855
2005	61,429,537	61,394,933	780,800
2006	61,661,640	61,644,695	749,559
2007	61,672,723	61,658,778	818,600
2008	63,842,647	63,831,902	724,242
2009	67,689,537	67,679,080	790,494
2010	71,330,572	71,322,080	813,206
2011	74,953,622	74,932,372	839,599
2012	77,166,643	77,135,657	848,215
2013†	50,383,104		813,451
2014	35,170,047		-

*NDI data available through 2013.

†Only 28 states submitted MAX data in 2013 and 17 submitted MAX data in 2014 due to transition from MSIS to TMSIS.

Conditions Segment

The CCW contains variables used to indicate the presence of treatment for common or chronic conditions using claims-based algorithms (as a proxy for evidence of the presence of a condition). This information is present in the MESF Conditions segment for all people within the MAX PS, regardless of whether the person has any of the conditions. There is a row of data for each identifier within each state; that is, if there was a beneficiary identifier (BENE_ID) for the enrollee – there is a record for each state where the BENE_ID was found; approximately 91% of records have a BENE_ID. Similarly, if a BENE_ID was not available, there is a row of data for each MSIS identifier (MSIS_ID) for each state. For the Conditions segment, CCW performed a small amount of aggregation of records in the PS file; if the same person was enrolled in Medicaid in the same state for the year and had more than one MSIS_ID, there will be only one record in the Conditions segment.

> *Note:* Even though there is a record for each person by state combination (e.g., BENE_ID and STATE_CD or if the record does not have a BENE_ID, the MSIS_ID and STATE_CD), the information contained within the Conditions segment is person-level (i.e., variables identify whether the person was receiving treatment for the condition, regardless of state). This means that if there is more than one row of data for the enrollee, this information is duplicated on each row of the Conditions segment.
>
> This data structure allows for examination of condition prevalence at either the state or national level. Investigators are encouraged to use caution to avoid double-counting these records when calculating national condition statistics. We provide some SAS® code examples for common analytic scenarios in Appendix B.

The MESF Conditions segment includes the 27 CCW Medicare common and chronic conditions which were developed by CMS using a multi-stage process (listed in Table 2 below). Initially, ResDAC used CMS and other national data sources to identify candidate conditions which could

be coded using claims-based algorithms. Next, extensive literature reviews were conducted to gather code sets for each candidate condition. Finally, CMS engaged other federal agencies in a series of conversations to vet the proposed category definitions.

The Conditions segment also contains a wide variety of Other Chronic or Potentially Disabling Conditions, with algorithms developed by the CMS Medicare-Medicaid Coordination Office specifically to enhance research of the Medicare-Medicaid population (listed in Table 3 below). There are mental health and tobacco use conditions, as well as conditions which are often related to physical and intellectual disability and developmental disorders.

The condition variables consider administrative claims criteria, FFS versus managed care coverage criteria, and specified look-back time periods from which to observe a pattern of care. The claims criteria consider variations of the following:
- ICD-9, Current Procedural Terminology (CPT) 4, or Healthcare Common Procedure Coding System (HCPCS) codes
- Claim type(s) and count(s)
- Date(s) of service (e.g., claim thru dates at least one day apart)

The coverage criteria consider variations of Medicaid FFS and managed care coverage; similarly, for people who were dually enrolled in Medicaid and Medicare, they consider variations in Medicare Parts A, B, and HMO coverage. The specified time periods, or *reference time periods*, consider the length of time during which the clinical and coverage criteria are considered. The condition algorithms use calendar years as the basis for the reference years (e.g., 2010 data files with a variable for a condition algorithm with two-year reference period includes services between 01/01/09-12/31/2010). To be classified as meeting the coverage criteria, the person had to be covered for all 24 months of the reference period – or if the person died during this time frame, the person needed to have coverage all months alive (i.e., if the condition had a reference time period of 1 year, then the person had to have 12 months of FFS coverage), and have FFS coverage which would allow for observation of the claims of interest. For example, for the Medicaid-oriented variables for a condition using a one-year reference time period, the value "coverage met" means that the person had Medicaid FFS for all 12 months of the year (or all months enrolled during the year). If the value indicates coverage is not met, then the person had 11 or fewer months of FFS coverage, or no FFS coverage at all during the year.

The values for each of the variables indicate whether the person received services during the time frame to designate treatment for the condition (i.e., based on the FFS administrative claims pattern, the person likely is being treated for the condition – or not). The reference time period is the look-back period during which the other criteria must be met. It is possible for a person to meet the claims criteria for a given year and not the next year.

The following are valid values for each of the condition variables:
0. Neither claims nor coverage met
1. Claims met, coverage not met
2. Claims not met, coverage met

3. Claims and coverage met

Three variables are present for each of the conditions to represent different payer perspectives; however two of the perspectives should only be considered for use if the person is dually enrolled in Medicare. The three perspectives for the conditions variables are Medicaid, Medicare, and combined (Medicaid & Medicare dually enrolled). Each perspective relies on a different universe of claims and enrollment data:

- Medicaid perspective - uses only Medicaid enrollment and claims. The source data is exclusively the MAX data files. These variables use the *_MEDICAID naming convention (e.g., for atrial fibrillation, the condition is called AFIB_MEDICAID)
- Medicare perspective - uses only Medicare coverage and claims. The source data is CCW Medicare enrollment and claims data. These variables use the *_MEDICARE naming convention (e.g., AFIB_MEDICARE)
- Dually enrolled/combined perspective - uses a combination of Medicare and Medicaid enrollment and claims. The source is both MAX and Medicare enrollment and claims data. These variables use the *_COMBINED naming convention (e.g., AFIB_COMBINED)

Within a payer perspective, each of the condition variables has the same denominator; for example, for all conditions that use the Medicaid perspective, the denominator for each of the *_MEDICIAD variables is the same. If a person is not also eligible for Medicare, then all values for the Medicare versions of the conditions (naming convention *_MEDICARE) will be null/missing; similarly, the *_COMBINED values will be missing.

Note: An easy way to determine whether your study population includes people who were dually enrolled in Medicare & Medicaid is by using the DM_SOURCE variable; where D=person had a Medicare enrollment record during the year, as verified by CCW.

Investigators are cautioned that the COMBINED and MEDICARE versions of the Conditions variables only have meaning for people who are dually enrolled. An additional caution: if you wish to use the COMBINED variables, which consider both Medicare and Medicaid, it is important to note that "coverage met" means all 12 months of BOTH Medicare and Medicaid FFS.

Table 2. CCW 27 conditions and variable names

27 CCW conditions	Medicaid-only perspective	Combined perspective *(applies to the dually enrolled population)*	Medicare-only perspective *(applies to the dually enrolled population)*
Acquired Hypothyroidism	HYPTHY _MEDICAID	HYPTHY_COMBINED	HYPTHY _MEDICARE
Acute Myocardial Infarction	AMI_MEDICAID	AMI_COMBINED	AMI_MEDICARE
Alzheimer's Disease	ALZ_MEDICAID	ALZ_COMBINED	ALZ_MEDICARE
Alzheimer's or Related Dementia	ALZRDSD_MEDICAID	ALZRDSD_COMBINED	ALZRDSD_MEDICARE
Anemia	ANEMIA_MEDICAID	ANEMIA_COMBINED	ANEMIA_MEDICARE
Asthma	ASTHMA_MEDICAID	ASTHMA_COMBINED	ASTHMA_MEDICARE

27 CCW conditions	Medicaid-only perspective	Combined perspective *(applies to the dually enrolled population)*	Medicare-only perspective *(applies to the dually enrolled population)*
Atrial Fibrillation	AFIB_MEDICAID	AFIB_COMBINED	AFIB_MEDICARE
Benign Prostatic Hyperplasia	HYPPLA_MEDICAID	HYPPLA_COMBINED	HYPPLA_MEDICARE
Cataract	CAT_MEDICAID	CAT_COMBINED	CAT_MEDICARE
Chronic Kidney Disease (CKD)	CKD_MEDICAID	CKD_COMBINED	CKD_MEDICARE
Chronic Obstructive Pulmonary Disease (COPD)	COPD_MEDICAID	COPD_COMBINED	COPD_MEDICARE
Heart Failure	CHF_MEDICAID	CHF_COMBINED	CHF_MEDICARE
Depression	DEPR_MEDICAID	DEPR_COMBINED	DEPR_MEDICARE
Diabetes	DIAB_MEDICAID	DIAB_COMBINED	DIAB_MEDICARE
Glaucoma	GLCM_MEDICAID	GLCM_COMBINED	GLCM_MEDICARE
Hip/Pelvic Fracture	HFRAC_MEDICAID	HFRAC_COMBINED	HIRAC_MEDICARE
Hyperlipidemia	HYPLIP_MEDICAID	HYPLIP_COMBINED	HYPLIP_MEDICARE
Hypertension	HYPTEN_MEDICAID	HYPTEN_COMBINED	HYPTEN_MEDICARE
Ischemic Heart Disease	IHD_MEDICAID	IHD_COMBINED	IHD_MEDICARE
Osteoporosis	OST_MEDICAID	OST_COMBINED	OST_MEDICARE
Rheumatoid/ Osteoarthritis	RAOA_MEDICAID	RAOA_COMBINED	RAOA_MEDICARE
Stroke/ TIA	STRK_MEDICAID	STRK_COMBINED	STRK_MEDICARE
Breast Cancer	BRC_MEDICAID	BRC_COMBINED	BRC_MEDICARE
Colorectal Cancer	CRC_MEDICAID	CRC_COMBINED	CRC_MEDICARE
Lung Cancer	LNGC_MEDICAID	LNGC_COMBINED	LNGC_MEDICARE
Prostate Cancer	PRC_MEDICAID	PRC_COMBINED	PRC_MEDICARE
Endometrial Cancer	ENDC_MEDICAID	ENDC_COMBINED	ENDC_MEDICARE

Table 3. Behavioral, Mental Health, and disability-related conditions and variable names

Mental Health, Tobacco, Conditions

27 CCW conditions	Medicaid-only perspective	Combined perspective *(applies to the dually enrolled population)*	Medicare-only perspective *(applies to the dually enrolled population)*
Conduct Disorders and Hyperkinetic Syndrome	ACP_MEDICAID	ACP_COMBINED	ACP_MEDICARE
Anxiety Disorders	ANXI_MEDICAID	ANXI_COMBINED	ANXI_MEDICARE
Bipolar Disorder	BIPL_MEDICAID	BIPL_COMBINED	BIPL_MEDICARE
Major Depression	DEPSN_MEDICAID	DEPSN_COMBINED	DEPSN_MEDICARE
Personality Disorders	PSDS_MEDICAID	PSDS_COMBINED	PSDS_MEDICARE

27 CCW conditions	Medicaid-only perspective	Combined perspective *(applies to the dually enrolled population)*	Medicare-only perspective *(applies to the dually enrolled population)*
Post-traumatic Stress Disorder (PTSD)	PTRA_MEDICAID	PTRA_COMBINED	PTRA_MEDICARE
Schizophrenia	SCHI_MEDICAID	SCHI_COMBINED	SCHI_MEDICARE
Schizophrenia and Other Psychotic Disorders	SCHIOT_MEDICAID	SCHIOT_COMBINED	SCHIOT_MEDICARE
Tobacco Use Disorders	TOBA_MEDICAID	TOBA_COMBINED	TOBA_MEDICARE

Conditions Related to Intellectual, Developmental, and Physical Disability

27 CCW conditions	Medicaid-only perspective	Combined perspective *(applies to the dually enrolled population)*	Medicare-only perspective *(applies to the dually enrolled population)*
Autism Spectrum Disorder	AUTISM_MEDICAID	AUTISM_COMBINED	AUTISM_MEDICARE
Cerebral Palsy	CERPAL_MEDICAID	CERPAL_COMBINED	CERPAL_MEDICARE
Cystic Fibrosis and Other Metabolic Developmental Disorders	CYSFIB_MEDICAID	CYSFIB_COMBINED	CYSFIB_MEDICARE
Epilepsy	EPILEP_MEDICAID	EPILEP_COMBINED	EPILEP_MEDICARE
Intellectual Disabilities and Related Conditions	INTDIS_MEDICAID	INTDIS_COMBINED	INTDIS_MEDICARE
Learning Disabilities and Other Developmental Delays	LEADIS_MEDICAID	LEADIS_COMBINED	LEADIS_MEDICARE
Mobility Impairments	MOBIMP_MEDICAID	MOBIMP_COMBINED	MOBIMP_MEDICARE
Multiple Sclerosis and Transverse Myelitis	MULSCL_MEDICAID	MULSCL_COMBINED	MULSCL_MEDICARE
Muscular Dystrophy	MUSDYS_MEDICAID	MUSDYS_COMBINED	MUSDYS_MEDICARE
Sensory – Deafness and Hearing Impairment	HEARIMP_MEDICAID	HEARIMP_COMBINED	HEARIMP_MEDICARE
Sensory – Blindness and Visual Impairment	VISUAL_MEDICAID	VISUAL_COMBINED	VISUAL_MEDICARE
Spina Bifida and Other Congenital Anomalies of the Nervous System	SPIBIF_MEDICAID	SPIBIF_COMBINED	SPIBIF_MEDICARE
Spinal Cord Injury	SPIINJ_MEDICAID	SPIINJ_COMBINED	SPIINJ_MEDICARE
Traumatic Brain Injury and Nonpsychotic Mental Disorders due to Brain Damage	BRAINJ_MEDICAID	BRAINJ_COMBINED	BRAINJ_MEDICARE

27 CCW conditions	Medicaid-only perspective	Combined perspective *(applies to the dually enrolled population)*	Medicare-only perspective *(applies to the dually enrolled population)*
Other Developmental Delays	OTHDEL_MEDICAID	OTHDEL_COMBINED	OTHDEL_MEDICARE

Other Chronic Physical and Behavioral Health Conditions

27 CCW conditions	Medicaid-only perspective	Combined perspective *(applies to the dually enrolled population)*	Medicare-only perspective *(applies to the dually enrolled population)*
Alcohol Use Disorders	ALCO_MEDICAID	ALCO_COMBINED	ALCO_MEDICARE
Drug Use Disorders	DRUG_MEDICAID	DRUG_COMBINED	DRUG_MEDICARE
Fibromyalgia, Chronic Pain and Chronic Fatigue	FIBRO_MEDICAID	FIBRO_COMBINED	FIBRO_ MEDICARE
HIV/AIDS*	HIVAIDS_MEDICAID	HIVAIDS_COMBINED	HIVAIDS_ MEDICARE
Leukemia and Lymphoma	LEUKLYMPH_ MEDICAID	LEUKLYMPH_COMBINED	LEUKLYMPH_ MEDICARE
Liver Disease, Cirrhosis, and Other Liver Conditions (excluding Hepatitis)	LIVER_ MEDICAID	LIVER_COMBINED	LIVER_ MEDICARE
Hepatitis (Any Viral)	HEPVIRAL_MEDICAID	HEPVIRAL_COMBINED	HEPVIRAL_MEDICARE
Hepatitis A	HEPA_MEDICAID	HEPA_COMBINED	HEPA_MEDICARE
Hepatitis B - acute	HEPB_ACT_MEDICAID	HEPB_ACT_COMBINED	HEPB_ACT_MEDICARE
Hepatitis B - chronic	HEPB_CHR_MEDICAID	HEPB_CHR_COMBINED	HEPB_CHR_MEDICARE
Hepatitis C - acute	HEPC_ACT_MEDIC AID	HEPC_ACT_COMBINED	HEPC_ACT_MEDICARE
Hepatitis C - chronic	HEPC_CHR_MEDIC AID	HEPC_CHR_COMBINED	HEPC_CHR_MEDICARE
Hepatitis C - unspecified	HEPC_UNS_MEDICAID	HEPC_UNS_COMBINED	HEPC_UNS_MEDICARE
Hepatitis D	HEPD_MEDICAID	HEPD_COMBINED	HEPD_MEDICARE
Hepatitis E	HEPE_MEDICAID	HEPE_COMBINED	HEPE_MEDICARE
Migraine and Other Chronic Headache	MIGRAINE_ MEDICAID	MIGRAINE_COMBINED	MIGRAINE_ MEDICARE
Obesity	OBESITY_ MEDICAID	OBESITY_COMBINED	OBESITY_ MEDICARE
Opioid Use Disorder (Any of the three below)	OUD_ANY_COMBINED	OUD_ANY MEDICAID	OUD_ANY MEDICARE
OUD Diagnosis	OUD_DX_COMBINED	OUD_DX MEDICAID	OUD_DX MEDICARE
OUD Hospitalization or ED visit	OUD_HOSP_COMBINED	OUD_ HOSP MEDICAID	OUD_ HOSP MEDICARE
OUD Medication Assisted Therapy (MAT)	OUD_MAT_COMBINED	OUD_ MAT MEDICAID	OUD_ MAT MEDICARE

27 CCW conditions	Medicaid-only perspective	Combined perspective *(applies to the dually enrolled population)*	Medicare-only perspective *(applies to the dually enrolled population)*
Pressure Ulcers and Chronic Ulcer	ULCERS_ MEDICAID	ULCERS_COMBINED	ULCERS_ MEDICARE
Peripheral Vascular Disease	PVD_MEDICAID	PVD_COMBINED	PVD_MEDICARE
Sickle Cell Disease (SCD)	SCD_MEDICAID	SCD_COMBINED	SCD_MEDICARE

Analytic Considerations

Data file users should consider carefully which payer perspective variables to use for analyses, particularly if examining conditions for people who may have some Medicare-Medicaid enrollment (i.e., DM_SOURCE=D). For example, if a beneficiary meets the criteria for a Medicaid condition and the variable has a value of 3 = "Claims and coverage met" using the "*_MEDICAID" variables, but the person has very little Medicare coverage during the reference time period, this same person will be classified as a "1 = Claims met, coverage not met" for the same condition using the "*_COMBINED" variables. This difference in values is due to the person's coverage criteria in the COMBINED files indicating that the person did not have BOTH Medicare and Medicaid coverage for all months in the reference period. Appendix B of this document demonstrates two examples for how to use the Chronic/Conditions File to achieve different study objectives (see Code Example 1 and Code Example 2).

Investigators are cautioned that claims data are generally not available if the person is enrolled in managed care; this is always true for Medicare, and often true for Medicaid, with some state-by-state variation in the completeness of managed care encounter claims. Since the conditions algorithms use diagnoses on claims, there is not an opportunity to determine whether the managed care enrollees have been treated for the condition(s) of interest. The condition variables are designed to examine patterns of services – which serve as a proxy for indicating the person likely is receiving treatment for the condition.

Investigators are encouraged to determine whether restrictions to the study population may be appropriate (e.g., limiting the analysis of conditions to beneficiaries with FFS coverage), which can be accomplished by using the enrollment criteria variables within the MAX PS. A CCW Technical Guidance paper which describes some important considerations when using the CCW conditions for calculating Medicare population statistics is available on the CCW website, under the "Analytic Guidance" tab (http://www.ccwdata.org/analytic-guidance/index.htm). Many of the same analytical steps would apply to MAX Conditions File users; however particular cohort criteria would be selected from the MAX PS file.

Proposed Analytic Steps

Several SAS code examples are shared in Appendix B to demonstrate various analytic perspectives which may be addressed with the MESF Conditions segment. In general, the major analytic steps we propose are:

A. Start with the MAX PS File
    1) Identify your population subset of interest from the MAX PS File – considering whether to examine conditions only for people with fee-for-service (FFS) coverage, for whom conditions determinations may be made. For data processing efficiency, we also recommend keeping only the PS variables necessary for your study since the file is wide.
    2) Subset the population into two groups - those with and without BENE_IDs, since the unique key for linking these records to the MESF Conditions file is different. For records without a BENE_ID, the unique key is MSIS_ID and STATE_CD.
    3) Determine whether you will want state-specific totals or national totals, since the methodology for aggregating the records is different. Then, roll up these 2 denominator files to either the person by state-level (i.e., for state totals) or the unique person-level (i.e., for national totals).

B. Process the MES Conditions file
    4) Using the MESF Conditions file, decide which particular conditions are of interest and also which payer perspective(s) are appropriate for your population. For data processing efficiency, we recommend keeping only the condition variables necessary for your study since the file is wide.
    5) Divide the MESF Conditions file into two groups – those with and without BENE_IDs.
    6) Aggregate the records to accommodate either state-level or national-level data.

C. Merge your denominators with your CC files (i.e., merge the 2 BENE_ID files together, then the 2 files without BENE_IDs – by using the MSIS_ID and STATE_CD), keeping only the condition records for those in your denominator. Then combine the resulting files into an overall population/condition file.

D. Summarize conditions data at the state or national level.

There may be cases where there is a BENE_ID in the MAX PS file without a corresponding record in the MESF Conditions segment. This is due to differences in when the data files were produced. The condition files are static; at the point in time when the condition file was produced we may not have had sufficient information to use for assigning a BENE_ID to a particular record. However, the MAX PS file is extracted from the CCW database, where BENE_IDs are updated on a quarterly basis. Therefore, we may have acquired additional information to use for assigning a BENE_ID to MAX PS records; in CCW, once we can assign a BENE_ID, we do. If you encounter these unmatched records in the Conditions file – use the MSIS_ID and STATE_CD to merge the files rather than the BENE_ID.

NDI Segment

The variables in these segments of the file contain official death certificate data purchased by CCW from the National Death Index (NDI) through the Centers for Disease Control and Prevention (CDC). The original source of this information is state vital statistics offices, which record information from death certificates (see CDC documentation re: the NDI at: http://www.cdc.gov/nchs/data_access/ndi/about_ndi.htm). The NDI segment contains the cause of death.

This segment contains only the subset of enrollees who had a date of death in the MAX PS file; or if the person was dually enrolled in Medicare, a date of death in Medicare or Social Security Administration records. These data are available in an annual file for decedents from 1999 – 2013.

Note: You can tell whether the DOD was from Medicaid or Medicare sources by using the DM_SOURCE variable; when the value=D the person was dually enrolled in Medicare, and DOD the Medicare data was the original source.

Researchers wishing to obtain this NDI segment must obtain an additional approval beyond the CMS DUA. ResDAC will assist you with this process. The variables in the NDI segment, and the source data used to populate the variables, are displayed in Table 4.

Table 4. NDI Segment variables

Source	Variable Name	Label
CCW	BENE_ID	CCW Beneficiary identifier
CCW	MSIS_ID	MSIS identifier (from MAX PS record)
CCW	STATE_CD	Enrollee State Code (from MAX PS record)
CCW	DEATH_DT	Date of death in administrative data
CCW	DM_SOURCE	Source of Record: (D)ual or (M)edicaid Only
CCW	NDI_MATCH	Death confirmed through NDI match
NDI	NDI_DOD	Date of death according to NDI death certificate
NDI	DEATH_CERT_NUM	NDI Death Certificate Number
NDI	NDI_STATE_DEATH_CD	NDI State of Death (2-digit numeric)
NDI	ICD_CODE	ICD-10 Code for cause of death
NDI	ICD_TITLE	ICD-10 Title\label for cause of death code
NDI	ICD_CODE_113	113 ICD-10 Recodes (Cause of Death and enterocolitis due to Clostridium Difficile)
NDI	ICD_CODE_358	358 ICD-10 Recodes
NDI	ICD_CODE_130	130 ICD-10 Recodes (select causes of infant death)
NDI	RECORD_COND_1 (thru 20)	NDI Record Axis Cause of Death - Condition 1 (for 1999-2006 there were up to 8 variables, for 2007-2013 there are 20 variables, sequentially numbered)
NDI	ENTITY_COND_1 (thru 20)	NDI Entity Axis Cause of Death - Condition 1 (for 1999-2006 there were up to 8 variables, for 2007-2013 there are 20 variables, sequentially numbered)

When CCW obtained the NDI data from NCHS, we identified decedents and then NCHS searched for a corresponding death certificate record. The NCHS matching to death certificate data works best when particular identifiers are available for linkage to the population; however, some Medicaid decedents (i.e., those with a DOD in the MAX PS), we did not have adequate personal identifiers (e.g., some states did not record Social Security Numbers). The NDI_MATCH variable in the NDI segment provides information regarding whether NCHS returned an NDI record for the decedent.

Additional information regarding the cause of death recodes can be found at:
https://www.cdc.gov/nchs/data/dvs/Part9InstructionManual2011.pdf

Chapter 3. Inpatient Hospital (IP) Files

The IP file contains complete stay records for Medicaid enrollees who used inpatient hospital services. The IP file also includes:

- Admission and discharge dates
- Up to nine The International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnostic codes
- A principal procedure code and up to five additional procedure codes along with a procedure code modifier to indicate the coding system used for the procedure codes
- A code to indicate if the hospital stay was for a maternal or newborn delivery
- Information on the Diagnosis Related Group (DRG), as applicable
- For UB-92/UB-04, includes up to 23 revenue code groups
- Medicaid payment amounts for fee-for-service records; encounter records may also be present
- Third-party payments for these services, as collected by the Medicaid agency, for each IP service
- For dual enrollee IP services, Medicaid paid Medicare copayment and deductible amounts.
- A count of the number of covered inpatient days
- Provider number

Chapter 4. Long Term Care (LT) Files

The LT file includes institutional long-term care (LTC) records for services provided by four types of long-term care facilities: 1) mental hospitals for the aged; 2) inpatient psychiatric facilities for persons under age 21; 3) intermediate care facilities for the mentally retarded (ICF/MR); and, 4) nursing facilities (NF). These records do not include procedure codes. Note that other community-based LTC services (e.g., many home-based and personal care services) are included in the OT file. Important types of information in the LT file include:

- An indicator for the type of institutional LTC service (as listed above)
- Admission date, beginning and end dates of service
- Patient status (at discharge or end of service date on the claim)
- Up to five ICD-9-CM diagnosis codes. Users should note that these may be diagnoses at the time of LTC facility admission and are not necessarily current diagnoses
- Medicaid payment amounts for fee-for-service records
- Third-party payments for these services, as collected by the Medicaid agency, for each LT service
- A count of the Medicaid covered days
- Provider number

Chapter 5. Prescription Drug (RX) Files

This file contains prescribed drugs, over-the-counter drugs and other items dispensed by a free-standing pharmacy. However, there are important exceptions:

- Drugs provided during an inpatient hospital stay are included in the hospital record reported in the IP file. Detail is not available for those drugs in IP records.

- For most states, drugs provided to persons in LTC facilities are included in this file. But there are a small number of states that include (or "bundle") these drugs into their LTC facility payment rate. Detail is not available for drugs that are bundled this way. At last report (for 1999), three states: New York, South Dakota, and Delaware were bundling drugs into their LTC facility payment rate.

- Injectable drugs that are administered by a health professional are included in the OT file. They are identified as procedure (service) codes with a leading character of "J".

- Beginning in 2006, coverage of prescription drugs was available through Medicare Part D for those with dual eligibility (i.e., coverage through both Medicaid and Medicare programs). More information on the Medicare Part D prescription drug event (PDE) data can be found at (http://www.ccwdata.org). People who qualify for both Medicare and Medicaid (full-benefit dual eligibles) receive the Medicare Part D drug benefit and their RX utilization for Part D-covered drugs is found in the Medicare Part D Event data. Dual eligibles may have some drug information in the RX file, particularly for Medicare non-covered drugs (e.g., over the counter drugs).

Important types of information in the RX File include:

- Date of service (prescription fill date)
- New or refill indicator
- Medicaid payment amounts for fee-for-service records
- Third-party payments for these services, as collected by the Medicaid agency, for each prescription fill
- National Drug Code (NDC)
- Quantity and days supply
- Provider number

Changes beginning with the 2006 RX file affected two data elements (data elements 35, Medicare coinsurance payment amount and 36, Medicare deductable payment amount).

Chapter 6. Other Therapy (OT) Files

A wide variety of service types are included in the OT file, including physician and professional services, outpatient and clinic, DME, hospice, home health, lab/x-ray and others, which are indicated by the Type of Service codes on the claim.

The OT file contains two major types of records: 1) records for services delivered; and 2) payment records for premiums paid to three types of prepaid managed care plans (which can be identified using the type of service codes).

1) Capitated Payments for Health Maintenance Organization (HMO)/Health Insuring Organization (HIO)
2) Prepaid Health Plans (PHPs)
3) Primary Care Case Management (PCCM)

Because of the nature of billing for certain types of services, e.g., Healthcare Common Procedure Coding System (HCPCS) codes for ambulance services, multiple claims with the exact same enrollment/person information and service dates will appear within the OT file. These should not be viewed as errors or data anomalies. The presence of what appear to be duplicate records are not true duplicates, but rather distinct services or portions of a service provided (e.g., there are limits regarding the time/quantity of services which may be billed on one claim – if an ambulance ride goes beyond the specified distance limit, an additional claim/claims for the remainder of the distance may be submitted or home health services may be billed in 15 minute increments). DME and supplies (not including drugs) billed by non-pharmacy providers (and containing HCPCS or other state-specific procedure codes) are also reported in this file.

Important types of information in the OT file:

- Dates or dates of service
- Type of service
- Up to two ICD-9-CM diagnostic codes
- A procedure (service) code
- Medicaid payment amounts for fee-for-service records; encounter records
- Third-party payments for these services, as collected by the Medicaid agency, for each OT service
- For dual enrollee OT services, Medicaid paid Medicare copayment and deductible amounts
- Provider number
- Place of service

Chapter 7. Limitations/Considerations of MAX Files

The MAX data provide a fairly complete and accurate view of Medicaid and Medicaid expansion (M-CHIP enrollment) and payments and utilization for services provided under FFS arrangements. However, MAX encounter record utilization data (and source data from MSIS) may be incomplete for Medicaid enrollees in prepaid plans, such as managed care plans. States are required to submit encounter records for services provided to enrollees in prepaid plans. However, CMS analyses indicate that encounter reporting lacks the consistency and completeness to support many research activities.

A fairly large proportion of Medicaid enrolled people are dually enrolled in Medicare, as documented on the CMS Medicare-Medicaid Coordination website (at http://www.cms.gov/Medicare-Medicaid-Coordination/Medicare-and-Medicaid-Coordination/Medicare-Medicaid-Coordination-Office/Analytics.html). For these dual enrollees, it may be desirable to identify medial service use that is paid either by Medicare or Medicaid. The CMS Medicare Medicaid Coordination Office (MMCO) has worked with Buccaneer to create the Medicare-Medicaid Linked Enrollee Analytic Data Source (MMLEADS) data files. This suite of linked files is designed to allow for the examination of all Medicare and Medicaid enrollment and claims data for those who were dually enrolled in both programs. Both Medicare claims and MAX claims, as well as the corresponding beneficiary enrollment and demographic information, is available in annual calendar year data files. MMLEADS Files are available on an annual basis starting in 2006. The files are generally available around 3 years after the end of the observation year. For example, the 2009 files became available in MMLEADS in March 2013. A data dictionary and MMLEADS User Guide can be found on the CCW website (ccwdata.org).

Although a rare event, all five MAX file types may not necessarily be available for each year and/or state, e.g., there is not an IP, LT, or OT file for Maine for 2005 forward. In addition, completeness of data fields varies greatly by state. More information about data anomalies found in the MAX data can be found on the CMS website at (http://www.cms.gov/MedicaidDataSourcesGenInfo/07_MAXGeneralInformation.asp). Because states may make final changes when compiling MSIS data each year, anomalies may differ by individual states each year as well.

The anomalies report (found in the link immediately above) may not yet be available for the most current MAX data that is available for release. To make good use of the anomaly reports, a basic understanding of the Medicaid program is essential, as well as an understanding of the variables and values important for the research project. Once a study question or research hypothesis is formulated, it may be prudent to review the anomaly reports to ensure that the desired data will be obtained (i.e., no anomalies which might limit a researcher's choice of states, years or file types).

Chapter 8. Content, Format, and Encryption of CCW Output Files

This section describes the content and format of the CCW MAX data package (the CCW data that are delivered to researchers). The MAX record layouts for various years and file types may vary, so the number of files depends upon the specific data requested. Delivery files may be packaged to accommodate the appropriate record layout for the year/type of data requested as well as to meet the maximum file size limit. The "Read Me" document contained on the delivery media will describe the specific files delivered with the request. The table below demonstrates how the files may vary depending on the request. For example, a request for 2006 IP data for only one state would include 1 Read Me and 1 Executable (zip file); the zip file would contain 1 .SAS, 1 .csv, and 1 .fts file. A request for all MAX data (five file types) for one year for a small state would result in 1 Read Me and 1 Executable (zip) file: the zip file would contain five .SAS files, five .csv files, and five .fts files. Requests for highly populated states and/or for a number of years/file types may result in the delivery of several data files for a single file type.
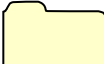
File Types on Delivery Media

Read Me Document .txt	Self-Decrypting Archive File .exe	SAS Read-in Program .sas	Data File .csv	File Transfer Summary .fts
One or more master Read Me file(s) depending on request	One or more files depending on years of data requested	By file type and year •	By type, state, and year of file	By file type and year of file

The files are delivered to the researcher in an organized format. File name/Descriptions for these items are detailed in the tables that follow.

File	File Description
readme_first_req000xxxx.txt	This is a text file that describes the files contained in the output package. File Name Example: readme_first_ req000xxxx.txt. The file provides information about the DUA, Request Number, and Year(s) of the data. This file also explains the three file types (.csv / .fts/ .sas) contained within the self-decrypting archive (.exe).

File	File Description
req000xxxx_MAX_2007.exe	This is the executable program that must be run to decrypt and uncompress the MAX data file. In this example, xxxx is the request number, and 2007 is the file year of the data. This executable includes v8 SAS read-in programs, the .csv file, and .fts file which contains the layout and record counts.

In addition to the specific data files the researcher requested, CCW includes a variety of resource files in the deliverable package. These files are described below.

File	Description
Decryption Instructions.pdf	This document contains instructions for decrypting/uncompressing the data files.
Shipping Package Insert.doc	Brief note to recipient.
File Record layouts	Folder contains variable listings for all CCW data file types. Also contains a sub-folder with corresponding data dictionaries.

Chapter 9. Decryption of Data Files

The encryption technique for files extracted from the CCW uses Pretty Good Privacy (PGP) Command Line 9.0 with the Self-Decrypting Archive (SDA) method. This method builds a compressed, encrypted, password protected file using a FIPS 140-1/140-2 approved AES256 cipher algorithm. The SDA is built on the CCW production server, downloaded to a desktop PC, and burned to a Compact Disc (CD), Digital Video Disc (DVD), or Universal Serial Bus (USB) external hard drive depending on the size of the files.

After the data media is shipped to the researcher, the password to decrypt the archive is sent to the researcher by electronic mail. Each researcher request will have a unique encryption. The password and the data media will never be packaged together. To decrypt the data files, the researcher will need to access the e-mail containing the decryption password. Detailed instructions for using this password are included with the data.

The CCW BENE_ID and encrypted MSIS_ID are specific to the CCW (not applicable to any other identification system or data source). All requested data are linked using the BENE_ID, and it is not available, data are linked using both the encrypted MSIS_ID and STATE_CD. The BENE_ID and MSIS_ID are encrypted using a contractor-developed cipher prior to delivery of data files to researchers. The cipher used is unique for each DUA and is determined at the time the data are requested. This key is then kept on file for future use, if requested by a researcher, and approved by CMS. A researcher may decide to stipulate in a new DUA that the data obtained must be linked to that obtained from a previous DUA. CMS will then evaluate and approve or disapprove the request. If approved, the data obtained from the CCW will be encrypted using the same cipher as the previous DUA allowing data from both requests to be linked.

Chapter 10. Further Assistance with CCW Data

The data dictionary for the CCW MAX files is available on the CCW website (https://www.ccwdata.org/web/guest/data-dictionaries).

The Research Data Assistance Center (ResDAC) offers free assistance to researchers using MAX data for research. The ResDAC web site provides links to descriptions of the CMS data available, request procedures, supporting documentation, such as record layouts and SAS input statements, workshops on how to use MAX data, and other helpful resources. Statistical summaries and links to additional information are also available on this site. Visit the ResDAC web site at (http://www.resdac.org) for additional information.

ResDAC is a CMS contractor and requests for assistance in the application, obtaining, or using the CCW data should first be submitted to ResDAC. Researchers can reach ResDAC by phone at 1-888-973-7322, e-mail at resdac@umn.edu, or online at (http://www.resdac.org/)

In the event that a ResDAC technical advisor is not able to answer the question, the technical advisor will direct the researcher to the appropriate person. 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 contact ResDAC for further assistance.

<div align="center">

Prepared by: HealthAPT, LLC
www.ccwdata.org
Email: CMSdata@gdit.com
Phone: 1-866-766-1915

</div>

Appendix A. List of Acronyms and Abbreviations

Acronym	Definition
BENE_ID	A unique beneficiary identifier assigned to all EDB and Medicaid eligible records as they are added to the CCW
CCW	Chronic Conditions Data Warehouse; constructed by Buccaneer, A General Dynamics Company under contract with CMS
CMS	Centers for Medicare & Medicaid Services; part of the U.S. Department of Health and Human Services which administers the Medicare, Medicaid and CHIP programs
DME	Durable Medical Equipment. Found in RX and OT files
DOD	Date of death
DRG	Diagnosis Related Group; a prospective payment system; each DRG has a payment weight assigned to it, based on the average resources used to treat Medicare patients in that DRG
DUA	Data Use Agreement delineates the confidentiality requirements of CMS regarding the Privacy Act and data release policies
EDB	Enrollment Database. The EDB is the CMS repository of enrollment and entitlement data for persons who are or have ever been enrolled in Medicare. On a monthly basis the repository is "unloaded" into a flat file (referred to as the Unloaded EDB), which then is loaded to the CCW
FQHC	Federally Qualified Health Centers are designated by the DHHS Health Resources and Services Administration (HRSA) to serve populations in medically underserved areas. FQHCs are community-based organizations that provide comprehensive primary care and preventive care, including health, oral, and mental health/substance abuse services to persons of all ages, regardless of their ability to pay
HCBCS	Home and Community-Based Services
HCBCA	Home and Community-Based Care for Aged and Elderly Disabled
HCPCS	Healthcare Common Procedure Coding System, a standardized coding system used by CMS to describe the specific items and services provided in the delivery of health care
IHS	Indian Health Service. An Operating Division within the U.S Department of Health and Human Services that is responsible for providing medical and public health services to members of federally recognized Tribes and Alaska Natives
IP	Inpatient Data
LT	Long Term Care Data
MAX	Medicaid Analytic eXtract data files compiled from CMS Medicaid data
MDS	Minimum Data Set – assessment data required by CMS to be submitted by nursing facilities
MESF	Medicaid Enrollee Supplemental File
MMLEADS	Medicare-Medicaid Linked Enrollee Analytic Data Source

Acronym	Definition
MSIS	Medicaid Statistical Information System, compilation of state information regarding enrollment, service utilization, and payment
NDC	National Drug Code(s), indicate the drug product which was received; may be found in RX and OT files
NDI	National Death Index
NF	Nursing Facilities; claims may be found in LT files
OASIS	Outcome and Assessment Information Set – assessment data required by CMS to be submitted by home health agencies
OT	Other Data
OTC	Over-the-Counter drugs; medications which do not require a prescription
PS	Person Summary Data
RHC	Rural Health Clinics are clinics located in a rural, medically under-served area in the United States that has a separate reimbursement structure from the standard medical office under the Medicare and Medicaid programs. RHCs were established by the Rural Health Clinics Act (P.L. 95-210), (Section 1905 of the Social Security Act), and receive the RHC designation by the Health Resources and Services Administration (HRSA)
RX	Prescription Drug Data
SAF	Standard Analytic File; term used to describe a CMS data product which meets certain standardized specifications, which vary by type of data file (e.g., Medicare Part D SAF versus Medicare inpatient SAF)
SSA	Social Security Administration
SSN	Social Security Number; an identifier created by the SSA. The original intent was to track benefits eligibility; therefore historically some persons may not have received a SSN. Presently, SSNs are assigned at birth and used as a universal identifier

Appendix B. SAS® Coding Examples for MESF-Conditions File

These coding examples use the 2009 MAX data files. There are times when we show a record count for the data files to illustrate the impact of the processing step; the counts you obtain may differ, depending on the number of unduplicated BENE_IDs in the MAX files at the time your data files are extracted from CCW. Our examples are in SAS®; however you can adapt the code to use any statistical software you prefer.

Code example 1: Calculate state-specific Medicaid prevalence of select conditions.

Begin by using the MAX PS to identify your study population. To create state-specific totals we will want to keep a record for each person within each state; that is, we do not want to remove or collapse records if a person is enrolled in more than one state (as we will do for national totals in).

> **STEP 1. Determine records that will link using BENE_ID.** The following $SAS^®$ code assumes that you are subsetting the MAX PS into two smaller files – one with records that have BENE_IDs and one without BENE_IDs (called *ps2009_BI* and *ps2009_noBI*, respectively). We summarize monthly variables to enable determination of which enrollees have fee-for-service (FFS) coverage.

2009 PS file N=67,689,538

```
data ps2009_BI ps2009_noBI;
      set a.max_ps_2009 (keep= BENE_ID MSIS_ID STATE_CD EL_DOD
      EL_PHP_TYPE: MAX_ELG_CD:);

/*assign a monthly FFS indicator based on Monthly Medicaid Eligibility
where uniform eligibility is not '00' and monthly PHP type where PHP
is not '01'(comprehensive managed care),  '05'(LTC managed care), or
'06'(PACE)*/

      array php1 (*) $ EL_PHP_TYPE_1_1-EL_PHP_TYPE_1_12;
      array php2 (*) $ EL_PHP_TYPE_2_1-EL_PHP_TYPE_2_12;
      array php3 (*) $ EL_PHP_TYPE_3_1-EL_PHP_TYPE_3_12;
      array php4 (*) $ EL_PHP_TYPE_4_1-EL_PHP_TYPE_4_12;
      array elg (*) $ MAX_ELG_CD:;
      array ffs (*) ffs1-ffs12;
      do i=1 to 12;
            if elg(i)~= '00' and php1(i) not in ('01','05','06') and
            php2(i) not in ('01','05','06') and
            php3(i) not in ('01','05','06') and
            php4(i) not in ('01','05','06') then ffs(i)=1;
            else if elg(i)~= '00' and (php1(i) in ('01','05','06') or
            php2(i) in ('01','05','06') or
            php3(i) in ('01','05','06') or
            php4(i) in ('01','05','06')) then ffs(i)=0;
            else ffs(i)=.;
```

```
      end;

      /*output a file for enrollees with BENE_IDs*/;

      if BENE_ID~= ' ' then output ps2009_BI;

      /*output a file for those without BENE_IDs*/;

      if BENE_ID='  ' then output ps2009_noBI;

run;
```

> **STEP 2. Aggregate PS file to BENE_ID by state level**. For your file with BENE_IDs (called ***ps2009_BI***),
> roll up records to the BENE_ID by state level (variable called STATE_CD). This step collapses records for
> those with more than one record in a state and retains the record with full fee for service status since 1 indicates
> FFS and 0 indicates non-FFS.
>
> For your file without BENE_IDs (called ***PS2009_noBI***), no state-level roll-up is possible. Keep all rows as in
> the source PS file. There is no STEP 2 for these records.

```
proc sort data=ps2009_BI;
     by BENE_ID STATE_CD;
run;


/*summarize data at the state-level by keeping the maximum of
the FFS indicator*/;

proc means data=ps2009_BI noprint;
     by BENE_ID STATE_CD;
     id EL_DOD;
     output out=bene_state(drop=_freq_ _type_) max(ffs: )= ;
run;


/*For the population without BENE_IDs, the MSIS_ID and STATE_CD
is the unique key; these records in are already aggregated at
that level (i.e., in the source MAX PS)*/
```

> **STEP 3. Link PS and Conditions file.** Work with the MES_CC File to subset into two smaller files – one with
> records that have BENE_IDs and one without BENE_IDs (called ***all_cc_bene*** and ***all_cc_nobene***, respectively).
>
> For your file with BENE_IDs (called ***ps2009_BI***), roll up records to the BENE_ID by state-level (variable called
> STATE_CD). This step collapses records for those with more than one record in a state and retains the record
> with the highest values for the condition variables (i.e., those which represent status using full FFS coverage).

```
data all_cc_bene  all_cc_nobene;            ┌─────────────────────────┐
     set m.MES_CC_2009;                     │ all_cc_bene N=62,107,891 │
     if BENE_ID~= ' ' then output all_cc_bene; └─────────────────────────┘
     if BENE_ID='  ' then output all_cc_nobene;
```

```
run;
```

For the 2009 MES_CC (N=67,679,185), 91.7% of the records (62,107,891) have a BENE_ID.

```
STEP 3A.   Work with subset who have BENE_IDs;
Re-calculate rate using a bene by state aggregation, keeping the
highest values for the condition variables (e.g., a 2 rather
than 0 and a 3 rather than 1). This will result in an increased
chance of identifying an individual with the condition and FFS.

proc means data=all_cc_bene noprint;
     by BENE_ID STATE_CD;
     id MSIS_ID;
     output out=bene_cc_state (drop=_freq_ _type_) max(ACP_:
     AFIB_: ALZRDSD_: ALZ_: AMI_: ANEMIA_:
     ANXI_: ASTHMA_: AUTISM_: BIPL_: BRAINJ_: BRC_:
     CAT_: CERPAL_: CHF_:CKD_: COPD_: CRC_: CYSFIB_:
     DEPR_: DEPSN_: DIAB: ENDC_: EPILEP_: GLCM_:
     HEARIM_: HFRAC_: HYPLIP_: HYPPLA_: HYPTEN_: HYPTHY_:
     IHD_: INTDIS_: LEADIS_: LNGC_: MOBIMP_: MULSCL_:
     MUSDYS_: OST_: OTHDEL_: PRC_: PSDS_: PTRA_:
     RAOA_: SCHIOT_: SCHI_: SPIBIF_: SPIINJ_: STRK_:
     TOBA_: VISUAL_:)=;
run;
```

bene_cc_state N=62,087,340

```
/*Link the BENE/STATE level file for those with a BENE_ID to the
MES_CC file using BENE_ID and STATE_CD as the linkage key*/

/*Retaining the STATE_CD allows for state level rates based on a
unique count of enrollees within a state*/

data bene_state_cc;
     merge bene_state (in=f) bene_cc_state;
     by BENE_ID STATE_CD;

if f;
run;


STEP 3B.   Work with subset who do not have bene_IDs;

/*For this subset there is not a need for a PROC MEANS statement
as in 3A since we can not aggregate the data below the MSIS_ID
by STATE_CD level*/

proc sort data=ps2009_noBI;
     by MSIS_ID STATE_CD;
```

```
run;

proc sort data=all_cc_nobene;
      by MSIS_ID STATE_CD;
run;

data no_bene_state_cc;
      merge ps2009_noBI (keep=MSIS_ID STATE_CD ffs:  in=f)
      all_cc_nobene;
      by MSIS_ID STATE_CD;

      if f;
run;
```

> **STEP 4.** Append the analytic files for those with and without BENE_IDs (i.e., combine *bene_state_cc* and *no_bene_state_cc*, respectively); determine which people have full FFS coverage (variable called ffs2009). This output file (called *all_state*) will be used to calculate state-level condition rates.

```
data a.all_state;
      set bene_state_cc no_bene_state_cc;

/*Determine those who for FFS Medicaid for all months alive
in 2009 for denominator inclusion*/

      if year(EL_DOD)=2009 and month(EL_DOD)=sum(of ffs: )
      then ffs2009=1;
      else if year(EL_DOD)~=2009 and sum(of ffs:)=12 then
      ffs2009=1;
      else ffs2009=0;
run;
```

For everyone in the MAX PS file the *_MEDICAID set of conditions variables is populated - which consider Medicaid FFS coverage and claims criteria. The *_COMBINED and *_MEDICARE perspectives for the condition variables will only be populated for the subset of enrollees who were also dually enrolled in Medicare (i.e., where DM_SOURCE="D").

For this coding example, we look across the condition variables at all 3 payer perspectives (the _MEDICAID, _COMBINED, and _MEDICARE variables) to determine whether there is evidence of treatment for the condition (where the value includes the 1s or 3s for each of the perspectives); you may wish to look only at _MEDICAID versions of the variable if you have identified a Medicaid FFS population.

> **STEP 5.** Calculate state-level condition rates. We demonstrate this using s subset of the conditions (atrial fibrillation, acute myocardial infarction, and heart failure). Note that we give an example that includes all conditions in Code example 2.

```
data heart_state_ffs (drop=ACP_: AFIB_: ALZRDSD_: ALZ_: AMI_: ANEMIA_:
     ANXI_: ASTHMA_: AUTISM_: BIPL_: BRAINJ_: BRC_:
     CAT_: CERPAL_: CHF_:CKD_: COPD_: CRC_: CYSFIB_:
     DEPR_: DEPSN_: DIAB_: ENDC_: EPILEP_: GLCM_:
     HEARIM_: HFRAC_: HYPLIP_: HYPPLA_: HYPTEN_: HYPTHY_:
     IHD_: INTDIS_: LEADIS_: LNGC_: MOBIMP_: MULSCL_:
     MUSDYS_: OST_: OTHDEL_: PRC_: PSDS_: PTRA_:
     RAOA_: SCHIOT_: SCHI_: SPIBIF_: SPIINJ_: STRK_:
     TOBA_: VISUAL_:);
     set a.all_state;

/*consider the values from all 3 payer perspectives*/

     if AFIB_MEDICAID in (1,3) or AFIB_COMBINED in (1,3) or
     AFIB_MEDICARE in (1,3) then afib=1;
     else afib=0;
     if AMI_MEDICAID in (1,3) or AMI_COMBINED in (1,3) or AMI_MEDICARE
     in (1,3) then ami=1;
     else ami=0;
     if CHF_MEDICAID in (1,3) or CHF_COMBINED in (1,3) or CHF_MEDICARE
     in (1,3) then chf=1;
     else chf=0;

/*subset file – keeping only those who were FFS Medicaid for all
months alive in 2009*/

if ffs2009=1;
run;


/* calculate state-specific rates*/

proc sort data=heart_state_ffs;
     by STATE_CD;
run;

proc freq data=heart_state_ffs;
     tables STATE_CD*(afib AMI CHF)/ nocol nopercent;
run;
```

There are many possible ways to tabulate the conditions data. In place of the PROC FREQ you may wish to use something like the following:

```
/*Aggregate numerator condition counts to state level*/
proc means data=heart_state_ffs noprint;
     by STATE_CD;
     output out=heart_state_cond(drop=_type_) sum(afib ami chf)=;
```

```
run;


/*Calculate State level rates*/
data heart_state_cond (drop= afib ami chf i);
     set heart_state_cond;
     array cond (*)  afib  ami chf ;
     array cond_rate (*) afib_pct ami_pct chf_pct ;
     do i=1 to dim(cond);
          cond_rate(i)=round(cond(i)/_freq_ *100,.1);
     end;
run;
```

A sample of the output from this final portion of the code example appears in Table 5, below.

Table 5. State-specific condition rates from code example 1

State	enrollee count	acp_pct	afib_pct	alzrdsd _pct	alz_pct	ami_pct	anemia_ pct	anxi_pct	asthma _pct
AK	69,920	4.6	0.8	2.2	0.6	0.1	5.5	7	3.2
AL	598,294	5.6	1.3	4	2.1	0.2	9.9	5.3	6.1
AR	524,946	8.1	1	3.4	1.6	0.2	5.9	5.2	3.7
AZ	157,690	0.9	0.4	0.8	0.2	0.1	3.9	1.6	2
CA	3,124,177	1.5	1.3	3.6	1.5	0.2	10.8	4	3.1
CO	295,381	2.5	1	4.1	1.8	0.1	6.5	4.9	4.8
CT	117,505	2.7	6.4	19.8	9.5	0.8	25.2	16.3	9.1
DC	41,827	4.2	1.8	8.2	3.9	0.6	20.2	5.5	7.2
DE	32,994	2.4	3.8	11.4	5.3	0.6	19.4	8.6	6.3
FL	1,068,017	3.2	2.1	7.7	4	0.3	16	6.9	5.8

Code example 2: Calculate national Medicaid prevalence of select conditions.

To fulfill this analytic objective, you begin processing the MAX files as detailed in Code Example 1, by partitioning the source data into separate files for those with and without BENE_IDs. Then, instead of aggregating the data by beneficiary and state, you will need to create a unique person-level file (i.e., to get rid of multiple rows of data for people enrolled in Medicaid in more than one state during the year – or in the same state during the year). Therefore, this STEP 2 is different than (but parallel to) Code Example 1.

STEP 2. Aggregate PS file to BENE_ID regardless of state. For your file with BENE_IDs (called ***ps2009_BI***, which was created in Code example 1), roll up records to the BENE_ID level.   This step collapses records for those with more than one record (either within a state or in a different state) and retains the record with full fee for service status since 1 indicates FFS and 0 indicates non-FFS.

For your file without BENE_IDs (called ***PS2009_noBI***), no state-level roll-up is possible. Keep all rows as in the source PS file; there is no STEP 2 for these records.

```
/*roll-up to BENE_ID for national analyses keeping the Maximum
of FFS indicator*/

/*summarize data at the national level*/;

proc means data=ps2009_BI noprint;
     by BENE_ID;
     id EL_DOD;
     output out=bene(drop=_freq_ _type_) max(ffs: )= ;
run;
```

```
Step 2B.  Work with subset who do not have BENE_IDs;

/*For this subset there is not a need for a PROC MEANS statement
as in 2A since we can not aggregate the data below the MSIS_ID
by STATE_CD level*/
```

Then, begin processing the MESF Conditions segment as explained in STEP 3 of Code Example 1 - by partitioning the source data into separate files for those with and without BENE_IDs. You may wish to limit it to the conditions of interest before attempting to join it with your population finder file.

> **STEP 3. Link PS and Conditions file** Work with the MES_CC File to subset into two smaller files – one with records that have BENE_IDs and one without BENE_IDs (called ***all_cc_bene*** and ***all_cc_nobene***, respectively; exactly as in Code Example 1). For your file with BENE_IDs (called ***ps2009_BI***), roll up records to the BENE_ID by state-level (variable called STATE_CD). This step collapses records for those with more than one record in a state and retains the record with the highest value for the condition giving preference for those with identification of FFS.
>
> In the DATA step, you may wish to keep only variables for the conditions of interest.

```
STEP 3A.  Work with subset who have BENE_IDs;
Re-calculate bene condition rates nationally keeping the highest
values for the condition variables (e.g., a 2 rather than 0 and
a 3 rather than 1). This will result in an increased chance of
identifying an individual with the condition and with FFS.

proc means data=all_cc_bene noprint;
     by bene_id;
     id msis_id;
     output out=bene_cc (drop=_freq_ _type_) max(acp_: afib_:
     alzrdsd_: alz_: ami_: anemia_:
     anxi_: asthma_: autism_: bipl_: brainj_: brc_:
     cat_: cerpal_: chf_:ckd_: copd_: crc_: cysfib_:
     depr_: depsn_: diab: endc_: epilep_: glcm_:
     hearim_: hfrac_: hyplip_: hyppla_: hypten_: hypthy_:
     ihd_: intdis_: leadis_: lngc_: mobimp_: mulscl_:
```

```
        musdys_: ost_: othdel_: prc_: psds_: ptra_:
        raoa_: schiot_: schi_: spibif_: spiinj_: strk_:
        toba_: visual_:)=;
run;


/*Aggregate the PS information that is in the working file
called bene that was created in step 2, and condition
information that is in the working file called bene_cc to an
enrollee level.
This will allow for national rates based on a unique count of
enrollees nationally*/

data bene_ntl_cc;
     merge bene bene_cc;
     by bene_id;
run;



STEP 3B.  Work with subset who do not have BENE_IDs;

/*You will need to use the no_bene_state_cc file that was
created in Step 3 from Code example 1*/
```

STEP 4. Append the analytic files for those with and without BENE_IDS (i.e., combine ***bene_ntl_cc*** and ***no_bene_state_cc***, respectively); determine which people have full FFS coverage (variable called ffs2009). This output file (called ***ntl_state***) will be used to calculate national condition rates.

```
data a.ntl_state;
     set bene_ntl_cc no_bene_state_cc;

/*Determine who has FFS Medicaid for all months alive in 2009
for denominator inclusion*/

     if year(EL_DOD)=2009 and month(EL_DOD)=sum(of ffs: ) then
     ffs2009=1;
     else if year(EL_DOD)~=2009 and sum(of ffs:)=12 then
     ffs2009=1;
     else ffs2009=0;
run;
```

For this coding example, we look across the condition variables at all 3 payer perspectives (the _MEDICAID, _COMBINED, and _MEDICARE variables) to determine whether there is evidence of treatment for the condition (where the value includes the 1s or 3s for each of the perspectives); you may wish to look only at _MEDICAID versions of the variable if you have identified a Medicaid FFS population.

> **STEP 5.** Calculate national condition rates. We demonstrate this using all of the condition variables. Note that we give a simpler example that includes a subset of conditions in Code Example 1.

```
data ntl_ffs (drop=acp_: afib_: alzrdsd_: alz_: ami_: anemia_:
      anxi_: asthma_: autism_: bipl_: brainj_: brc_:
      cat_: cerpal_: chf_:ckd_: copd_: crc_: cysfib_:
      depr_: depsn_: diab_: endc_: epilep_: glcm_:
      hearim_: hfrac_: hyplip_: hyppla_: hypten_: hypthy_:
      ihd_: intdis_: leadis_: lngc_: mobimp_: mulscl_:
      musdys_: ost_: othdel_: prc_: psds_: ptra_:
      raoa_: schiot_: schi_: spibif_: spiinj_: strk_:
      toba_: visual_:);

      set a.ntl_state;

/*look at all 3 payment perspectives for the each of the conditions*/

      if acp_medicaid in (1,3) or acp_combined in (1,3) or acp_medicare
      in (1,3) then acp=1;
      else acp=0;
      if afib_medicaid in (1,3) or afib_combined in (1,3) or
      afib_medicare in (1,3) then afib=1;
      else afib=0;
      if alzrdsd_medicaid in (1,3) or alzrdsd_combined in (1,3) or
      alzrdsd_medicare in (1,3) then alzrdsd=1;
      else alzrdsd=0;
      if alz_medicaid in (1,3) or alz_combined in (1,3) or alz_medicare
      in (1,3) then alz=1;
      else alz=0;
      if ami_medicaid in (1,3) or ami_combined in (1,3) or ami_medicare
      in (1,3) then ami=1;
      else ami=0;
      if anemia_medicaid in (1,3) or anemia_combined in (1,3) or
      anemia_medicare in (1,3) then anemia=1;
      else anemia=0;
      if anxi_medicaid in (1,3) or anxi_combined in (1,3) or
      anxi_medicare in (1,3) then anxi=1;
      else anxi=0;
      if asthma_medicaid in (1,3) or asthma_combined in (1,3) or
      asthma_medicare in (1,3) then asthma=1;
      else asthma=0;
      if autism_medicaid in (1,3) or autism_combined in (1,3) or
      autism_medicare in (1,3) then autism=1;
      else autism=0;
      if bipl_medicaid in (1,3) or bipl_combined in (1,3) or
      bipl_medicare in (1,3) then bipl=1;
      else bipl=0;
      if brainj_medicaid in (1,3) or brainj_combined in (1,3) or
      brainj_medicare in (1,3) then brainj=1;
      else brainj=0;
```

```
if brc_medicaid in (1,3) or brc_combined in (1,3) or brc_medicare
in (1,3) then brc=1;
else brc=0;
if cat_medicaid in (1,3) or cat_combined in (1,3) or cat_medicare
in (1,3) then cat=1;
else cat=0;
if cerpal_medicaid in (1,3) or cerpal_combined in (1,3) or
cerpal_medicare in (1,3) then cerpal=1;
else cerpal=0;
if chf_medicaid in (1,3) or chf_combined in (1,3) or chf_medicare
in (1,3) then chf=1;
else chf=0;
if ckd_medicaid in (1,3) or ckd_combined in (1,3) or ckd_medicare
in (1,3) then ckd=1;
else ckd=0;
if copd_medicaid in (1,3) or copd_combined in (1,3) or
copd_medicare in (1,3) then copd=1;
else copd=0;
if crc_medicaid in (1,3) or crc_combined in (1,3) or crc_medicare
in (1,3) then crc=1;
else crc=0;
if cysfib_medicaid in (1,3) or cysfib_combined in (1,3) or
cysfib_medicare in (1,3) then cysfib=1;
else cysfib=0;
if depr_medicaid in (1,3) or depr_combined in (1,3) or
depr_medicare in (1,3) then depr=1;
else depr=0;
if depsn_medicaid in (1,3) or depsn_combined in (1,3) or
depsn_medicare in (1,3) then depsn=1;
else depsn=0;
if diab_medicaid in (1,3) or diab_combined in (1,3) or
diab_medicare in (1,3) then diab=1;
else diab=0;
if endc_medicaid in (1,3) or endc_combined in (1,3) or
endc_medicare in (1,3) then endc=1;
else endc=0;
if epilep_medicaid in (1,3) or epilep_combined in (1,3) or
epilep_medicare in (1,3) then epilep=1;
else epilep=0;
if glcm_medicaid in (1,3) or glcm_combined in (1,3) or
glcm_medicare in (1,3) then glcm=1;
else glcm=0;
if hearim_medicaid in (1,3) or hearim_combined in (1,3) or
hearim_medicare in (1,3) then hearim=1;
else hearim=0;
if hfrac_medicaid in (1,3) or hfrac_combined in (1,3) or
hfrac_medicare in (1,3) then hfrac=1;
else hfrac=0;
if hyplip_medicaid in (1,3) or hyplip_combined in (1,3) or
hyplip_medicare in (1,3) then hyplip=1;
else hyplip=0;
```

```
if hyppla_medicaid in (1,3) or hyppla_combined in (1,3) or
hyppla_medicare in (1,3) then hyppla=1;
else hyppla=0;
if hypten_medicaid in (1,3) or hypten_combined in (1,3) or
hypten_medicare in (1,3) then hypten=1;
else hypten=0;
if hypthy_medicaid in (1,3) or hypthy_combined in (1,3) or
hypthy_medicare in (1,3) then hypthy=1;
else hypthy=0;
if ihd_medicaid in (1,3) or ihd_combined in (1,3) or ihd_medicare
in (1,3) then ihd=1;
else ihd=0;
if intdis_medicaid in (1,3) or intdis_combined in (1,3) or
intdis_medicare in (1,3) then intdis=1;
else intdis=0;
if leadis_medicaid in (1,3) or leadis_combined in (1,3) or
leadis_medicare in (1,3) then leadis=1;
else leadis=0;
if lngc_medicaid in (1,3) or lngc_combined in (1,3) or
lngc_medicare in (1,3) then lngc=1;
else lngc=0;
if mobimp_medicaid in (1,3) or mobimp_combined in (1,3) or
mobimp_medicare in (1,3) then mobimp=1;
else mobimp=0;
if mulscl_medicaid in (1,3) or mulscl_combined in (1,3) or
mulscl_medicare in (1,3) then mulscl=1;
else mulscl=0;
if musdys_medicaid in (1,3) or musdys_combined in (1,3) or
musdys_medicare in (1,3) then musdys=1;
else musdys=0;
if ost_medicaid in (1,3) or ost_combined in (1,3) or ost_medicare
in (1,3) then ost=1;
else ost=0;
if othdel_medicaid in (1,3) or othdel_combined in (1,3) or
othdel_medicare in (1,3) then othdel=1;
else othdel=0;
if prc_medicaid in (1,3) or prc_combined in (1,3) or prc_medicare
in (1,3) then prc=1;
else prc=0;
if psds_medicaid in (1,3) or psds_combined in (1,3) or
psds_medicare in (1,3) then psds=1;
else psds=0;
if ptra_medicaid in (1,3) or ptra_combined in (1,3) or
ptra_medicare in (1,3) then ptra=1;
else ptra=0;
if raoa_medicaid in (1,3) or raoa_combined in (1,3) or
raoa_medicare in (1,3) then raoa=1;
else raoa=0;
if schiot_medicaid in (1,3) or schiot_combined in (1,3) or
schiot_medicare in (1,3) then schiot=1;
else schiot=0;
```

```
        if schi_medicaid in (1,3) or schi_combined in (1,3) or
        schi_medicare in (1,3) then schi=1;
        else schi=0;
        if spibif_medicaid in (1,3) or spibif_combined in (1,3) or
        spibif_medicare in (1,3) then spibif=1;
        else spibif=0;
        if spiinj_medicaid in (1,3) or spiinj_combined in (1,3) or
        spiinj_medicare in (1,3) then spiinj=1;
        else spiinj=0;
        if strk_medicaid in (1,3) or strk_combined in (1,3) or
        strk_medicare in (1,3) then strk=1;
        else strk=0;
        if toba_medicaid in (1,3) or toba_combined in (1,3) or
        toba_medicare in (1,3) then toba=1;
        else toba=0;
        if visual_medicaid in (1,3) or visual_combined in (1,3) or
        visual_medicare in (1,3) then visual=1;
        else visual=0;

/*subset file – keeping only those who were FFS Medicaid for all
months alive in 2009*/

        if ffs2009=1;
run;


/*Aggregate numerator counts to the national level and calculate the
national rates*/

proc means data=ntl_ffs noprint;
        output out=ntl_cond(drop=_type_) sum(acp afib alzrdsd alz ami
        anemia anxi asthma autism bipl brainj brc  cat cerpal chf ckd
        copd crc cysfib depr depsn diab endc epilep glcm hearim hfrac
        hyplip hyppla hypten hypthy ihd intdis leadis lngc mobimp mulscl
        musdys ost othdel prc psds ptra raoa  schiot  schi  spibif
        spiinj  strk toba  visual )=;
run;

data ntl_cond (drop=acp afib alzrdsd alz ami anemia anxi asthma autism
        bipl brainj brc cat cerpal chf ckd copd crc cysfib depr depsn
        diab endc epilep glcm hearim hfrac hyplip hyppla hypten hypthy
        ihd intdis leadis lngc mobimp mulscl musdys ost othdel prc psds
        ptra raoa  schiot  schi  spibif  spiinj  strk  toba  visual i);
        set ntl_cond;

/*Array to identify the count of individuals with each specific
condition. This is used in the do loop with the array below to
calculate percentages based on the population count*/
        array cond (*) acp afib alzrdsd alz ami anemia
        anxi asthma autism bipl brainj brc
        cat cerpal chf ckd copd crc cysfib
        depr depsn diab endc epilep glcm
```

```
      hearim hfrac hyplip hyppla hypten hypthy
      ihd intdis leadis lngc mobimp mulscl
      musdys ost othdel prc psds ptra
      raoa  schiot  schi  spibif  spiinj  strk
      toba  visual;

      array cond_rate (*) acp_pct afib_pct alzrdsd_pct alz_pct ami_pct
      anemia_pct anxi_pct asthma_pct autism_pct bipl_pct brainj_pct
      brc_pct cat_pct cerpal_pct chf_pct ckd_pct copd_pct crc_pct
      cysfib_pct depr_pct depsn_pct diab_pct endc_pct epilep_pct
      glcm_pct hearim_pct hfrac_pct hyplip_pct hyppla_pct hypten_pct
      hypthy_pct ihd_pct intdis_pct leadis_pct lngc_pct mobimp_pct
      mulscl_pct musdys_pct ost_pct othdel_pct prc_pct psds_pct
      ptra_pct  raoa_pct  schiot_pct  schi_pct  spibif_pct  spiinj_pct
      strk_pct  toba_pct  visual_pct;

      /*Calculation of percentages based on the population count
      (_freq_) and the count of individuals with a specific condition*/

      do i=1 to dim(cond);
            cond_rate(i)=round(cond(i)/_freq_ *100,.1);
      end;
run;
```

A sample of the output from this final portion of the code example appears in Table 6, below.

Table 6. National condition rates from code example 2

	enrollee count	acp_pct	afib_pct	alzrdsd_pct	alz_pct	ami_pct	anemia_pct	anxi_pct	asthma_pct
NTL	**20,184,677**	4.6	1.8	5.9	2.7	0.3	11.8	7.2	4.6

If you wish to combine national and state-level rates into a single report, you may add the following code:

```
/*Append the state rates with the national rates for an overall
summary file*/
data condition_summary;
      format mdcd_state_cd $char3.;
      set ntl_cond state_cond;
      if mdcd_state_cd=' ' then mdcd_state_cd='NTL';
run;
```