Chronic Condition Data Warehouse

Your source for national CMS Medicare and Medicaid research data

CCW Medicare Administrative Data User Guide

August 2018

Version 3.5

This page intentionally left blank.

Overview

One of the goals of Section 723 of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 was to make Medicare data readily available to researchers who are studying chronic illness in the Medicare population. To support this effort, the Centers for Medicare & Medicaid Services (CMS) established the Chronic Conditions Data Warehouse (CCW).

The CCW contains CMS administrative data (from multiple data sources) linked by a unique identifier, allowing researchers to analyze information across the continuum of care. The CCW currently contains data from Medicare, Medicaid, and all assessments regardless of payer. The Medicaid data are available to researchers as the Medicaid Analytic eXtract (MAX) data files.

CMS Data	Types of Files
Medicare	• Master Beneficiary Summary File (aka Denominator) • Plan Characteristics File • Institutional and Non-institutional fee-for-service (FFS) claims • Encounter data • Part D drug event fills • Part D Characteristics files
Medicaid (MAX files)	• Person Summary (PS) • Medicaid Enrollee Supplemental File (MESF) Inpatient Hospital (IP) • Long Term Care (LT) • Prescription Drug (RX) • Other Services (OT)
Assessment	• Minimum Data Set (MDS) • Outcome and Assessment Information Set (OASIS) • Swing bed assessments • Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF/PAI)

CCW data are available upon request for a random 5% sample or for specific chronic condition cohorts. Data may also be requested for other cohort(s) of interest. The specific chronic condition categories were defined by CMS and its contractors. Much of the supporting documentation is available on the CCW website (e.g., chronic condition definitions, standard data dictionary files, etc.). See www.ccwdata.org.

The intended use of the CCW data is to identify areas for improving the quality of care provided to chronically ill Medicare beneficiaries, reduce program spending, and make current Medicare data more readily available to researchers studying chronic illness in the Medicare population. By predefining the chronic conditions, data extraction from the CCW is very efficient, allowing for data requests to be fulfilled in a timely and cost efficient manner.

The CCW Medicare administrative claims files are provided to academic researchers and certain government agencies, which have been approved under a Data Use Agreement (DUA) to obtain Medicare administrative data for research purposes. The CCW Medicare 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 data http://www.resdac.org/).

Medicare beneficiary Health Insurance Claim numbers (HICs) are removed from the data files delivered to researchers (unless otherwise specified/approved in the DUA). A unique CCW beneficiary identifier (BENE_ID) is included in each data file delivered as part of the output package (see Chapter 7 for details), thus allowing linkage of an individual's data across data sources/types. A separate file is provided for those requests requiring beneficiary identifiable data. If a researcher needs to obtain the HIC in order to link to outside data sources or extract claims not part of the CCW database, then the researcher will need to submit justification for this information in the study protocol and request identifiable variables.

The unique CCW beneficiary identifier 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 7 for details). Each research request employs a different encryption key for the beneficiary identifier field and the data files.

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

Table of Contents

List of Tables

Chapter 1. CCW Medicare Population

The CCW is populated with CMS Medicare administrative enrollment and claims data for all Medicare beneficiaries with coverage during a specified time period. The CCW data are available for services beginning January 1, 1999 through the most current year of data available.

CCW contains enrollment data for 100% of Medicare beneficiaries, and FFS claims from 1999 forward. Managed care encounter data is available for people enrolled in Medicare Advantage (MA) plans in 2015 (the data dictionary for this file is available on ccwdata.org). CCW also contains all Part D events (prescription drug fill records) from the inception of the Part D benefit in 2006, regardless of whether the beneficiary was enrolled in a managed care plan or a stand-alone prescription drug plan.

This rich data repository allows for tremendous flexibility in defining cohorts or populations of interest. CCW offers various sampling options such as: a random 5% sample, a CCW condition cohort, or investigator-defined sample. Finder files from populations previously used by researchers may be used for sample selection if approved by CMS. A random 5% sample or a smaller population subset is encouraged, due to the very large volume of Medicare data.

The 5% random sample consists of people who had a Medicare HIC number equal to the Claim Account Number (CAN) plus Beneficiary Identity Code (BIC) (HIC=CAN+BIC) where the last two digits of the CAN are in the set {05, 20, 45, 70, 95}; an "enhanced" 5% sample consists of those who were ever part of the 5% sample at any time, beginning with January 1, 1999 forward. The HIC number is assigned by the Social Security Administration (SSA) when a person becomes eligible for benefits, however the number may change over time, if a person's reason for entitlement changes. The CAN number is basically the policy number of the wage earner who is eligible for benefits – which means the CANs for spouses are joined. A marriage may cause a change in the HIC due to entitlement for benefits through the spouse. Two variables are designed to make it easy to identify the random 5% sample for a particular year (variable called SAMPLE_GROUP), and also to follow the 5% sample longitudinally even when a HIC change causes the person to drop out of the 5% at a later point in time (variable called ENHANCED_FIVE_PERCENT_FLAG). Note that the SAMPLE_GROUP variable can also be used to identify a 1% or 20% sample.

Chapter 2. Medicare Enrollment and Beneficiary-Level Data

The CCW has always disseminated files which include data regarding Medicare enrollment. Historical files include the CMS Denominator File, the CCW Beneficiary Summary File (BSF), the Beneficiary Annual Summary File (BASF), and the Master Beneficiary Summary File (MBSF) that used the CMS Enrollment Database (EDB) as its source. Currently, the preferred enrollment data file is the MBSF that uses the CMS Common Medicare Environment (CME) Database as its source. The CCW transition from the EDB to the CME database as the source for the MBSF enrollment and demographic information offers a number of advantages that will better meet researchers' evolving data needs. A description of the rationale and impact of this data conversion is contained in a CCW White Paper, "Medicare Enrollment: Impact of Conversion from EDB to CME".

Like the original MBSF, the current MBSF using the CME enrollment data file contains a constellation of enrollment and other person-level variables contained in "segments" – which are separate components of the file researchers may elect to purchase. A description of the variables contained in the MBSF can be found on the Data Dictionary page on the CCW website at http://www.ccwdata.org/web/guest/data-dictionaries.

The MBSF is created annually and contains demographic entitlement and enrollment data for beneficiaries who: 1) were part of the user-requested sample; 2) were documented as being alive for some part of the reference year; and, 3) were enrolled in the Medicare program during the file's reference year. *Reference year* refers specifically to the calendar year accounted for in the MBSF. For example, the 2014 MBSF covers the year 2014 - which is the *reference year*.

The current MBSF – Part A/B/C/D segment, also known as the Base Beneficiary Summary File, consists of variables that identify monthly Medicare Part A, B, C, and D enrollment status and other key demographic and coverage variables. Some key demographic, enrollment and coverage variables are illustrated in Table 1 through Table 4. All of the fields regarding beneficiary demographics or enrollment in the MBSF are derived from the CMS Common Medicare Environment (CME) Database (which is also the source for the CMS Enrollment Database – EDB). CCW updates information each month, and allows a full year of additions and updates after the end of the calendar year before finalizing the MBSF. For example, the 2014 data were finalized in December 2015 and then extracted to populate the MBSF.

The additional segments of MBSF are: 1) CCW Conditions, 2) Other Chronic or Potentially Disabling Conditions, and 3) Cost and Use. These three optional segments are described below and listed in Table 5.

A. Medicare Part A, B, C and D Enrollment Segment

Essential information for most study denominators will appear in this segment of the MBSF. It consists largely of beneficiary demographic and Medicare Part A, B, C and D coverage information. Table 1 below identifies some key demographic variables, and Table 2 identifies Medicare enrollment and coverage variables.

Table 1. Examples of Demographic Information

Variable name (long)	Variable description	Brief definition
STATE_CODE	SSA State Code	Beneficiary SSA state code
COUNTY_CD	SSA County Code	Beneficiary SSA county code
STATE_CNTY_FIPS_CD_{*MM*}	Monthly State and County FIPS Code (01-12)	Beneficiary FIPS state and county code (12 monthly fields)
ZIP_CD	Zip Code of Residence	Beneficiary ZIP code
AGE_AT_END_REF_YR	Age at End of Reference Year	Age at end of reference year
BENE_BIRTH_DT	Date of Birth (DOB)	DOB from Social Security Administration (SSA)
BENE_DEATH_DT	Date of Death (DOD)	DOD from SSA
SEX_IDENT_CD	Sex	Beneficiary sex
BENE_RACE_CD	Beneficiary Race Code	Beneficiary race code
RTI_RACE_CD	Research Triangle Institute (RTI) Race Code	RTI race code

Beneficiary state, county and ZIP code – The beneficiary geographic information comes from the beneficiary's mailing address, which is used to deliver cash benefits to the beneficiary (such as Social Security) or for other purposes (such as Medicare premium billing). It may not reflect the location where the beneficiary resides.

RTI Race Code– This variation on the race code which has historically been used by the Social Security Administration (SSA), classifies an additional group of beneficiaries as Hispanics or Asians. Using this enhanced classification algorithm, Hispanics and Asians include beneficiaries who *either* have an SSA race code which = Hispanic or Asian or a first name/last name which RTI (note: the contractor which created this field is RTI International - a trade name of Research Triangle Institute) has determined is likely to be Hispanic or Asian in origin. The current version of the RTI race code includes CCW enhancements designed to reduce the number of beneficiaries with missing information.

Table 2. Examples of Medicare Enrollment Information

Variable name (long)	Variable description	Brief definition
COVSTART	Medicare Coverage Start Date	The first historical date for Medicare coverage
ENTLMT_RSN_ORIG	Original Reason for Entitlement Code	Original reason for entitlement to Medicare

Variable name (long)	Variable description	Brief definition
ENTLMT_RSN_CURR	Current Reason for Entitlement Code	Current (year) reason for entitlement to Medicare
ESRD_IND	End Stage Renal Disease (ESRD) Indicator	Beneficiary entitled to end-stage renal disease (ESRD) benefits
MDCR_STATUS_CODE_{*MM*}	Medicare Status Code (01-12)	Current (monthly) reason for entitlement to Medicare (12 monthly fields)
MDCR_ENTLMT_BUYIN_IND_{*MM}*	Medicare Entitlement/Buy-In Indicator (01-12)	Monthly indicator of entitlement to Medicare A and B, as well as whether a state paid Medicare premiums (12 monthly fields)
DUAL_STUS_CD_{*MM*}	State Reported Dual Eligible Status Code (01-12)	Monthly indicator of dual eligibility status; where beneficiary is enrolled in both Medicaid and Medicare (12 monthly occurrences).
BENE_STATE_BUYIN_TOT_MONS	State Buy-In Coverage Months Count	Number of months of state buy-in for the beneficiary
DUAL_ELGBL_MONS	Months of Dual Eligibility	Number of months where the beneficiary had dual eligibility (DUAL_STUS_CD_XX not equal to '00' or '**')

B. Medicare Part C

Beneficiaries may elect to receive original fee-for-service (FFS) Medicare or, as an alternative, enroll in Medicare Part C (Medicare Advantage, or MA). Medicare Advantage Organizations (MAOs) are private managed care plans, such as Health Maintenance Organizations (HMOs), Preferred Provider Organizations (PPOs), Private Fee-for-service plans (PFFS), and Special Needs Plans (SNPs), which provide Medicare Part A and Part B services. Most MA plans include the Medicare Part D prescription drug benefit, and are referred to as MA-PD plans. Part C enrollment variables are displayed in Table 3.

Table 3. Part C Enrollment Variable Descriptions

Variable name (long)	Variable description	Brief definition
HMO_IND_{*MM*}	Health Maintenance Organization (HMO) Indicator (01-12)	Monthly indicator of whether the beneficiary was enrolled in a managed care plan, currently referred to as Medicare Advantage premiums (12 monthly fields)
BENE_HMO_CVRAGE_TOT_MONS	HMO Coverage Months Count	Number of months where the beneficiary had Medicare Advantage (HMO) coverage
PTC_CNTRCT_ID_{*MM}*	Contract ID (01-12)	The unique number CMS assigns to each contract that a Part C plan has with CMS (12 monthly occurrences).

Variable name (long)	Variable description	Brief definition
		The first character of the contract ID is a letter representing the type of plan, e.g., Managed Care Organizations, Regional Preferred Provider Organization (PPO), Regional PPO, PDP, Not Part D Enrolled, Employer Direct Plan (beginning in 2007) (12 monthly occurrences).
PTC_PBP_ID_{*MM*}	Plan Benefit Package ID (01-12)	The unique number CMS assigns to identify a specific Part C plan benefit package within a contract (12 monthly occurrences).
PTC_PLAN_TYPE_CD_{*MM*}	Plan Type Code	Monthly Part C Plan Type Code (12 monthly occurrences).

Starting with the 2015 benefit year, CCW has made available a Plan Characteristics suite of six files per year that contains detailed information regarding the Part C and Part D Plans selected by beneficiaries. Additional details regarding Plan Characteristics are in (Chapter 5). For 2015, Medicare Encounter Records are also available; additional details are in a separate user guide on the ccwdata.org website.

C. Medicare Part D

The Medicare prescription drug benefit, which is a voluntary benefit offered though the Medicare Part D program, is optional drug coverage beneficiaries may purchase through private plans. Coverage of prescription drugs through Medicare Part D began in 2006. The Part D enrollment data are available in the MBSF. Part D variables are displayed in Table 4.

Table 4 Part D Enrollment Variable Descriptions

Variable name (long)	Variable description	Brief definition
PTD_CNTRCT_ID_{*MM}*	Contract ID* (01-12)	The unique number CMS assigns to each contract that a Part D plan has with CMS (12 monthly occurrences). The first character of the contract ID is a letter representing the type of plan, e.g., Managed Care Organizations, Regional Preferred Provider Organization (PPO), Regional PPO, PDP, Not Part D Enrolled, Employer Direct Plan (beginning in 2007)
PTD_PBP _ID_{*MM}*	Plan Benefit Package ID* (01-12)	The unique number CMS assigns to identify a specific Part D plan benefit package within a contract (12 monthly occurrences).

Variable name (long)	Variable description	Brief definition
PTD_SGMT_ID_{*MM}*	Segment ID* (01-12)	The segment number CMS assigns to identify a segment or subdivision of a Part D plan benefit package within a contract (12 monthly occurrences).
CST_SHR_GRP_CD_{*MM*}	Cost Share Group (01-12)	Monthly indicator of beneficiary liability of cost-sharing. Includes values to indicate whether the beneficiary was deemed to be eligible or whether there was a subsidy (12 monthly occurrences).
RDS_IND_{*MM*}	Retiree Drug Subsidy Indicators (01-12)	Monthly indicator of whether employer should be subsidized for retired beneficiary (12 monthly occurrences).
PTD_PLAN_CVRG_MONS	Part D Plan Coverage Months	Total number of months of Part D plan coverage
RDS_CVRG_MONS	Retiree Drug Subsidy Months	Total number of months employer is entitled to Retiree Drug Subsidy for beneficiary

*Prior to the 2013 data release, CCW was required to encrypt these sensitive data fields in accordance with CMS privacy rules. We are no longer required to encrypt the plan identifiers.

D. Conditions Segments

There are two conditions segments: 1) the CCW conditions, and 2) the CMS Other Chronic or Potentially Disabling Conditions.

CCW Conditions

The CCW Chronic Condition segment of the MBSF contains the 27 CCW Medicare common and chronic conditions which were developed by CMS using a multi-stage process. 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 algorithms examine patterns of services, which serve as a proxy indicating that a beneficiary is likely receiving treatment for the condition.

The CCW Chronic Conditions segment contains variables which 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 for all beneficiaries included in the requested sample, regardless of whether the person has any of the conditions. This file includes the *yearly*, *mid-year*, and *ever* flags for each of the chronic conditions

described in Chapter 3. These three variables appear for each of the conditions, as illustrated with the acute myocardial infarction (AMI) example in Table 5 below.

Table 5. Examples of CCW Condition Variables

Variable name (long)	Brief definition
AMI	Acute myocardial infarction (AMI) end-of-year indicator
AMI_MID	Acute myocardial infarction mid-year indicator
AMI_EVER	First occurrence of acute myocardial infarction

Additional details regarding the condition variables appear in Chapter 3 of this document.

Other Chronic or Potentially Disabling Conditions

The Other Chronic or Potentially Disabling Conditions segment of the MBSF contains 15 mental health and substance use conditions, 15 developmental disorder and disability-related conditions, and 9 other chronic physical and behavioral health conditions which were developed by CMS specifically to enhance research of the Medicare-Medicaid dually enrolled population. These variables are similar in structure to the variables in the CCW Condition segment; there is an end of the year version of the variables and an "ever" version. The variable naming convention includes *_MEDICARE to distinguish these variables from conditions in the CCW Conditions segment (Table 6). Additional details are presented in Chapter 3.

Table 6. Examples of Other Condition Variables

Variable name (long)	Brief definition
AUTISM_MEDICARE	Autism spectrum disorders end-of-year indicator
AUTISM_MEDICARE_EVER	Autism spectrum disorders first ever occurrence date

E. Cost and Use Segment

This segment of the MBSF contains summarized patient-level utilization information, by care setting for the calendar year of the data file. It also includes Medicare and beneficiary payment information overall, and by setting. These cost and use summaries use Medicare Part A and Part B fee-for-service claims; therefore, there is not an opportunity to determine whether the managed care enrollees have received services, nor are we able to calculate the annual costs of these services. The exception is for Part D events – where we have both the cost and use information regardless of whether the beneficiary is enrolled in a stand-alone prescription drug plan (PDP) or in a Medicare Advantage plan with prescription drug coverage (MA-PD).

Variables are null (missing) if the beneficiary did not use a particular type of service during the year. For example, those without an acute hospitalization will have missing values for all variables associated with the acute care setting.

Settings. To better illustrate the cost and use summaries, the methodology for dividing claims into settings is described in Tables 7, Table 10, Table 11 & 12. We classify Medicare services using four major setting categories:

- Part A Institutional Claims – Claims from institutions or facilities which are generally covered on the Medicare Part A benefit;
- Part B Institutional Claims – Claims from institutions such as hospital outpatient facilities, which are generally covered on the Medicare Part B benefit;
- Part B Non-Institutional Claims – Claims from non-institutional providers such as providers/practitioners and durable medical equipment or prosthetic/orthotics providers; and
- Part D Event Data – Final transactional record for all Medicare Part D prescription drug events.

Table 7. Algorithms Used in Categorizing Part A Institutional Claims into Settings

Part A Values	Label	Algorithm	Medicare Payment Variable Name
ACUTE	Inpatient acute care hospital (and CAH)	NCH_CLM_TYPE_CD= 60,61 and PRVDR_NUM has 3rd digit =0 or 3rd and 4th digits = 13	ACUTE_MDCR_PMT*
OIP	Other inpatient hospital (children's, cancer, IPF, IRF, LTC hospital)	all other NCH_CLM_TYPE_CD= 60,61 (where 3rd digit of PRVDR_NUM is not 0 and 3rd and 4th digits are not 13)	OIP_MDCR_PMT*
SNF	Skilled Nursing Facility	NCH_CLM_TYPE_CD= 20, 30	SNF_MDCR_PMT
HH	Home Health	NCH_CLM_TYPE_CD= 10	HH_MDCR_PMT
HOS	Hospice	NCH_CLM_TYPE_CD= 50	HOS_MDCR_PMT

* Note that for these hospital settings this variable may not completely represent the total amount paid by Medicare. To calculate total Medicare payments you must add the pass-through-per-diem payments to this variable.

Cost or Payments. Three types of payment variables are present in the MBSF-CU:

1. Medicare payments - represent the annual amount Medicare paid for services on behalf of the beneficiary. The payments are aggregated for each beneficiary for the year (i.e., sum of all Part A CLM_PMT_AMT and Part B LINE_NCH_PMT_AMT and Part D CPP_AMT and LIS_AMT). For hospital settings, you must also add in the pass-through-per-diem payments in order to obtain the total Medicare payments (i.e., total Medicare payments for acute inpatient hospitalizations= ACUTE_MDCR_PMT + ACUTE_PERDIEM_PMT).

Additional information regarding Medicare payments is available in a CCW Technical Guidance document called "Getting Started with CMS Medicare Administrative Research Files," available on the CCW website (https://www.ccwdata.org/web/guest/technical-guidance-documentation).

2. Beneficiary payments – represent the aggregated beneficiary liability for cost-sharing, including coinsurance and deductible payments for the year. This includes the annual sum of all claims for one of the following (depending on the setting):

   - **Part A -** NCH_BENE_IP_DDCTBL_AMT and NCH_BENE_PTA_COINSRNC_LBLTY_AMT, all
   - **Part B Institutional** - REV_CNTR_CASH_DDCTBLE_AMT and REV_CNTR_WAGE_ADJSTD_COINS_AMT, and all
   - **Part B Non-Institutional** - B_DED, B_COINS, and
   - **Part D** - PTPAYAMT, OTHTROOP, LICS_AMT, and PLRO_AMT

3. Primary payer other than Medicare – represent the amount a primary payer (e.g., the VA or TRICARE) paid for services on behalf of the beneficiary. The payments are aggregated for each beneficiary for the year (i.e., sum of all Part A NCH_PRMRY_CLM_PD_AMT and all Part B CARR_CLM_PRMRY_PYR_PD_AMT).

These three types of payments are present for almost every service type. For example, the total annual Medicare payments for a beneficiary for SNF care can be found in the variable called SNF_MDCR_PMT, the corresponding beneficiary payments can be found in the SNF_BENE_PMT variable, and the other primary payer amounts can be found in the SNF_PRMRY_PMT variable. Two service types, the HOS and HH, do not have a beneficiary payment variable since the coinsurance and deductible amounts for these two settings were $0. The three types of payment variables that appear for nearly every service setting are illustrated in Table 10. Algorithm Used in Categorizing Part B Institutional Claims for the Part A service settings.

Table 8. Part A Service Settings and Corresponding Payment Variables

Service type	Medicare Payment	Beneficiary Payment	Primary Payer Amount
Inpatient acute care hospital (and CAH)	ACUTE_MDCR_PMT + ACUTE_PERDEIEM_PMT	ACUTE_BENE_PMT	ACUTE_PRMRY_PMT
Other inpatient hospital (IPF, Cancer hosp, children's hosp)	OIP_MDCR_PMT + OIP_PERDIEM_PMT	OIP_BENE_PMT	OIP_PRMRY_PMT
Skilled Nursing Facility	SNF_MDCR_PMT	SNF_BENE_PMT	SNF_PRMRY_PMT
Home Health	HH_MDCR_PMT		HH_PRMRY_PMT
Hospice	HOS_MDCR_PMT		HOS_PRMRY_PMT

The same three types of payment variables are present for the Part B Institutional Outpatient (variables called HOP_MDCR_PMT, HOP_BENE_PMT and HOP_PRMRY_PMT), and for each of the 11 Part B Non-institutional Outpatient settings that appear in Table 11 (all such variables follow the naming convention *_MDCR_PMT and *_BENE_PMT). For Part D drugs, in addition to the Medicare and beneficiary payment variables, there is a gross drug cost variable that is the annual sum of the total drug costs (i.e., TOTALCST accrued on behalf of the beneficiary for the year).

Utilization. Additional summary variables within the Cost and Use segment include many variables that identify service use in a granular fashion:

- **Stays** variables – count of hospital stays (unique admissions, which may span more than one facility) in the inpatient setting for a given year. A hospital stay is defined as a set of one or more consecutive inpatient claims where the beneficiary is only discharged on the most recent claim in the set. If a beneficiary is transferred to a different provider, the stay is continued even if there is a discharge date on the claim from which the beneficiary was transferred. The CLM_FROM_DT for the first claim associated with the stay must have been in the year of the data file, however, the algorithm allows for the CLM_THRU_DT to have occurred in January of the following year.
    - Acute stays – count of acute hospital stays during the year; variable called ACUTE_STAYS
    - Stays – in the OIP, SNF, and Hospice settings; variable called STAYS is the count of each type of stay during the year (OIP_STAYS, SNF_STAYS, HOS_STAYS)

- **Covered Days** – for acute, OIP, SNF, and Hospice settings; Medicare covered days (not all days for an institutional stay will be covered by Medicare). The covered days (variables in data file called ACUTE_COV_DAYS, OIP_COV_DAYS, SNF_COV_DAYS, and HOS_COV_DAYS) are calculated by summing the CLM_UTLZTN_DAY_CNT for the particular type of services for the year.

- **Readmissions** - count of hospital readmissions in the acute inpatient setting for a given year (variable called READMISSIONS). The CLM_FROM_DT for the original admission

must have been in the year of the data file, however, the algorithm allows for the readmission claim to have occurred in January of the following year.

A beneficiary is considered to be readmitted when she has an acute inpatient stay with a discharge status that indicates she has not expired (DSCHRG_STUS≠20) or left against medical advice (DSCHRG_STUS≠07) within 30 days of a previous acute inpatient stay that also has a discharge status that is not expired or left against medical advice. All beneficiaries without an ACUTE stay will have a missing value; beneficiaries with an ACUTE stay who do not have a subsequent readmission will have a value of 0 (zero).

- **Emergency Room Visits** - The MBSF-CU also identifies emergency room (ER) visit use. There are two setting scenarios we capture – ER visits where the patient became hospitalized at the same facility, or ER visits that were outpatient and the patient was not hospitalized at the same facility. Information for ER utilization is obtained by examining the revenue center records for the claims. ER revenue center codes were any of the following: 0450, 0451, 0452, 0456, or 0459. The inpatient ER visits (variable called IP_ER_VISITS) are a subset of ACUTE services; that is because the ER visit is not paid separately from the hospitalization. The hospital outpatient ER visits (variable called HOP_ER_VISITS) are a subset of HOP services. The sum of these two variables is the total ER use for the beneficiary for the year.

A listing of the variables that summarize utilization for Part A claims appears in Table 9.

Table 9. Part A Service Settings and Corresponding Utilization Variables

Service type	Medicare covered days	Stays	ER use	30-day readmissions
Inpatient acute care hospital (and CAH)	ACUTE_COV_DAYS	ACUTE_STAYS	IP_ER_VISITS	READMISSIONS
Other inpatient hospital*	OIP_COV_DAYS	OIP_STAYS		
Skilled Nursing Facility	SNF_COV_DAYS	SNF_STAYS		
Home Health		HH_VISITS		
Hospice	HOS_COV_DAYS	HOS_STAYS		

* Other hospitals include Inpatient psychiatric facility (IPF), cancer hospitals, and children's hospitals.

Table 10. Algorithm Used in Categorizing Part B Institutional Claims into a Setting

Part A Values	Label	Algorithm	Medicare Payment Variable Name
HOP	Hospital Outpatient	NCH_CLM_TYPE_CD= 40	HOP_MDCR_PMT

- **Visits**

- HH - the annual sum across all Part A claims for the number of home health visits on each claim (i.e., sum of CLM_HHA_TOT_VISIT_CNT); variable called HH_VISITS
- HOP - this variable is the count of unique revenue center dates (as a proxy for visits) in the hospital outpatient setting for a given year; variable called HOP_VISITS
- Emergency Room – two different variables 1) ER visits where the patient became hospitalized at the same facility (variable called IP_ER_VISITS), or 2) ER visits that were outpatient and the patient was not hospitalized at the same facility (variable called HOP_ER_VISITS), which are a subset of HOP services.

Table 11. Algorithms Used in Categorizing Part B Non-Institutional Claims into Settings

Non-Institutional Part B Values	Label	Algorithm: NCH_CLM_TYPE_CD= 71, 72, 81, 82 (BETOS codes)*	Medicare Payment Variable Name
ASC**	Ambulatory surgical center	LINE_CMS_TYPE_SRVC_CD="F" and NCH_CLM_TYPE_CD= 71 or 72	ASC_MDCR_PMT
ANES	Anesthesia	Anesthesia (P0) where CARR_LINE_MTUS_CD='2' and NCH_CLM_TYPE_CD= 71 or 72	ANES_MDCR_PMT
PTB_DRUG	Part B drug	Chemotherapy (O1D), Other Part B drug (O1E), Immunization (O1G), DME drug (D1G), Imaging drugs (I1E and I1F)	PTB_DRUG_MDCR_PMT
PHYS	Physician office	Physician office (M1A or M1B)	PHYS_MDCR_PMT
EM	Evaluation & management	Hospital (M2), Emergency room (M3), Home or nursing home visit (M4), Specialist (M5), and Consultation (M6)	EM_MDCR_PMT
DIALYS	Dialysis services	Dialysis services (P9)and NCH_CLM_TYPE_CD= 71 or 72	DIALYS_MDCR_PMT
OPROC	Other Procedures (not ANES or DIALYS)	Other – major procedure (P1), Major cardiac (P2), Major orthopedic (P3), Eye (P4), Ambulatory procedure (P5), Minor procedure (P6), Oncology procedure (P7), and Endoscopy (P8) - and NCH_CLM_TYPE_CD= 71 or 72	OPROC_MDCR_PMT
IMG	Imaging	Standard imaging (I1), Advanced imaging (I2), Echography (I3), and Imaging procedure (I4)	IMG_MDCR_PMT
TEST	Laboratory or test	Laboratory test (T1) and Other test (T2) - and NCH_CLM_TYPE_CD= 71 or 72	TEST_MDCR_PMT
DME	Durable medical equipment	DME supplies (D1A – D1E) and Orthotic devices (D1F)	DME_MDCR_PMT

Non-Institutional Part B Values	Label	Algorithm: NCH_CLM_TYPE_CD= 71, 72, 81, 82 (BETOS codes)*	Medicare Payment Variable Name
OTHC	Other Part B Carrier services	Ambulance (O1A), Chiropractic (O1B), Parenteral nutrition (O1C), Vision, hearing or speech services (O1F), and Other/unclassified Part B service (Y1, Y2, Z2, and missing)	OTHC_MDCR_PMT

* The first two or three digits of the identified BETOS codes appear in parentheses.
** The algorithms within this table are hierarchical – ASC must be identified first and OTHC must be last.

Additional summary variables within the Cost and Use segment related to the Part B Non-Institutional claims include:

- **Events** - An event is defined as each claim line item that contains the relevant service. The variables which summarize events count all relevant line items (i.e., line items corresponding with each type of service) for the beneficiary for the year. One such variable is PHYS_EVENTS (i.e., the number of occurrences [line items on the claim] for a physician face-to-face visit); there is an *_EVENTS variable corresponding to each Part B Non-Institutional setting in Table 11.

Table 12. Algorithm Used in Categorizing Part D Prescription Drug Events

Part D Values	Label	Algorithm	Total Payment Variable Name
PTD	Part D	PDE file	PTD_TOTAL_RX_CST

The Cost and Use segment also contains Part D summary cost and use; this information is available for all beneficiaries who have Part D coverage, even those in managed care plans (i.e., MA-PD plans):

- Part D Medicare (Part D plan) payment for filled prescriptions for covered drugs (variable called PTD_MDCR_PMT) – calculated as the sum of two CCW variables: the amount paid by the plan for Part D covered drugs (CVRD_D_PLAN_PD_AMT) and any low-income cost sharing (LIS) amount (LICS_AMT). Note: this variable does not include all costs to Medicare for the Part D benefit (also does not consider include any applicable rebate amounts or other discounts).
- Part D Beneficiary payment (variable called PTD_BENE_PMT; cost-sharing for filled prescriptions) - calculated as the sum of the CCW variables: patient pay amount (PTPAYAMT), other True Out-of-Pocket (TrOOP) amount (OTHTROOP), and patient liability reduction due to other payer (PLRO) amount (PLRO_AMT) for Part D drugs for a given year.
- Part D prescription events (PDE; variable called PTD_EVENTS) – count of all PDE IDs (i.e., unique prescription fill events) for the year.

- Part D fills (variable called PTD_FILL_CNT) - PDEs consist of highly variable days supply of the medication. This derived variable creates a standard 30 day supply of a filled Part D prescription, and counts this as a "fill". The Part D fill count does not indicate the number of different drugs the person is using, only the total months covered by a medication (e.g., if a patient is receiving a full year supply of a medication, whether this occurs in one transaction or 12 monthly transactions, the fill count = 12; if the patient is taking three such medications, the fill count=36).
- Part D total prescription cost (variable called PTD_TOTAL_RX_CST) - the gross drug cost (TOTALCST) of all Part D drugs for a given year. Note that the sum of the plan/Medicare and beneficiary share of the payments will not equal the total drug cost if there is a Part D low-income subsidy (LIS) or third party payer (e.g., the VA, or TRICARE).

F. National Death Index (NDI) Segment

This segment of the MBSF contains data purchased by CCW from the CDC. The original source of this information is state vital statistics offices, which record information from death certificates (see CDC documentation regarding the NDI at: https://www.cdc.gov/nchs/ndi/). **These data are available for decedents from 1999 – 2016.** Researchers wishing to obtain this NDI segment of the MBSF must obtain an additional approval beyond the CMS DUA. Data available in the NDI segment are displayed in Table 13.

Table 13. NDI Segment Variables

Variable name (long)	Variable description	Brief definition
NDI_DEATH_DT	NDI death date	NDI Date of Death
NDI_STATE_DEATH_CD	NDI state death code	NDI State of Death (SSA numeric code)
DEATH_CERT_NUM	NDI Death Certificate Number	NDI Death Certificate Number
ICD_CODE	ICD-10 code	ICD-10 cause of death code
ICD_TITLE	Label for ICD-10	ICD-10 Cause of Death Title (label)
ICD_CODE_358	358 ICD-10 Recodes	358 selected causes of death, ICD-10 Recodes
ICD_CODE_113	113 ICD-10 Recodes	113 selected ICD-10 Cause of Death and enterocolitis due to Clostridium Difficile
ICD_CODE_130	130 ICD-10 Recodes	130 selected ICD-10 causes of infant death
ENTITY_COND_1 (through 20)	Entity axis conditions	NDI Entity Axis Cause of Death – Condition (for 1999-2006 there were up to 8 variables, for 2007-2016 there are 20 variables, sequentially numbered)
RECORD_COND_1 (through 20)	Record axis conditions	NDI Record Axis Cause of Death – Condition (for 1999-2006 there were

		up to 8 variables, for 2007-2016 there are 20 variables, sequentially numbered)

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

Chapter 3. Condition Segments

The CCW makes it easy to study chronic diseases by incorporating twenty-seven (27) condition-specific variables which indicate treatment for the condition of interest. The CCW conditions are available in the MBSF CCW Conditions segment. In addition, we have included variables to identify 35 Other Chronic or Potentially Disabling Conditions which were created by CMS to enhance research of the Medicare-Medicaid dually enrolled population. Information regarding these conditions is available in the MBSF Other Chronic or Potentially Disabling Conditions segment. Medicare enrollment/eligibility and FFS claims data for each beneficiary are available within CCW (note: Medicare claims data are generally not available if the beneficiary is enrolled in managed care).

The predefined conditions use claims-based algorithms to indicate that treatment for a condition appears to have taken place; therefore, there is not an opportunity to determine whether the managed care enrollees have been treated for the condition(s) of interest. This limitation also applies, perhaps to a lesser extent, to newly-eligible Medicare beneficiaries who may have only a partial year of FFS coverage. 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. Medicare claims used the ninth version of the International Classification of Diseases (ICD-9) to classify all diagnoses, which identify the condition(s) for which a patient is receiving care, until September 2015. CMS switched to the next version of diagnosis codes, ICD-10-CM, for discharges and services on October 1, 2015. All CCW condition algorithms have been converted to use ICD-10 codes, for October 2015 forward.

It is important to note that the major objective of the chronic condition (CC) indicator variables is to allow for relevant clinical cohorts to be easily extracted from a very large database. The CCW condition definitions were intended to be somewhat broad, so that more researchers could request data extractions based on these definitions – then refine the specifications as needed to fit their own data needs. The condition definitions are not necessarily intended for investigators to use for calculating population statistics.

Investigators are encouraged to determine whether restrictions to the obtained CCW Condition segment and/or Other Chronic or Potentially Disabling Conditions segment of the MBSF should be made for their analyses. A CCW Technical Guidance paper which describes some important considerations when using the CCW conditions for calculating population statistics is available on the CCW website (http://www.ccwdata.org/web/guest/technical-guidance-documentation).

Researchers may request data files for cohorts based on *standard*, *modified standard*, or *custom* definitions. Note: 17 of the original 21 CCW condition algorithms were revised, and made available when 2010 data were delivered; these updated CCW conditions are available from 1999 forward.

Standard definitions include specific criteria for reference time periods, diagnosis and procedure codes, number/type of qualifying claims (e.g., must have two Carrier claims during reference time period), coverage (see CC category descriptions below), and exclusions. Researchers may request CCW data for any of the 27 predefined CCW chronic conditions as defined by CMS. These common chronic disease classifications are displayed in Table 14 (also, see http://www.ccwdata.org/web/guest/condition-categories). Cohorts may also be identified using the Other Chronic or Potentially Disabling Conditions displayed in Table 15.

Table 14. CCW Chronic Condition Classifications

Acquired Hypothyroidism*	Chronic Kidney Disease
Acute Myocardial Infarction	Chronic Obstructive Pulmonary Disease & Bronchiectasis
Alzheimer's Disease	Depression
Alzheimer's Disease, Related Disorders, or Senile Dementia	Diabetes
Anemia*	Glaucoma
Asthma*	Heart Failure
Atrial Fibrillation	Hip / Pelvic Fracture
Benign Prostatic Hyperplasia*	Hyperlipidemia*
Cancer, Colorectal	Hypertension*
Cancer, Endometrial	Ischemic Heart Disease
Cancer, Female/Male Breast	Osteoporosis
Cancer, Lung	Rheumatoid Arthritis / Osteoarthritis
Cancer, Prostate	Stroke / Transient Ischemic Attack
Cataract	

** The asterisk indicates the condition was not one of the original 21 conditions.*

Table 15. Other Chronic or Potentially Disabling Conditions

Alcohol Use Disorders Anxiety Disorders Autism Spectrum Disorder Bipolar Disorder Cerebral Palsy Conduct Disorders and Hyperkinetic Syndrome Cystic Fibrosis and Other Metabolic Developmental Disorders Depressive Disorders Drug Use Disorders Epilepsy Fibromyalgia, Chronic Pain and Chronic Fatigue HIV/AIDS Intellectual Disabilities and Related Conditions Learning Disabilities and Other Developmental Delays Leukemia and Lymphoma Liver Disease, Cirrhosis, and Other Liver Conditions (excluding Hepatitis) Migraine and Other Chronic Headache Mobility Impairments Multiple Sclerosis and Transverse Myelitis	Muscular Dystrophy Obesity Other Developmental Delays Opioid Use Disorder (OUD; overarching indicator) • Diagnosis and Procedure Basis for OUD • Opioid-Related Hospitalization or Emergency Department Visit • Use of Medication-Assisted Treatment (MAT) Peripheral Vascular Disease Personality Disorders Post-traumatic Stress Disorder (PTSD) Pressure Ulcers and Chronic Ulcer Sensory – Blindness and Visual Impairment Schizophrenia Schizophrenia and Other Psychotic Disorders Sensory-Deafness and Hearing Impairment Spina Bifida and Other Congenital Anomalies of the Nervous System Spinal Cord Injury Traumatic Brain Injury and Nonpsychotic Mental Disorders due to Brain Damage Tobacco Use Disorders Viral Hepatitis

Within each of the 27 CCW variables and 39 Other Condition variables, there are values which indicate whether the beneficiary received services during the time frame to designate treatment for the condition (i.e., based on the FFS administrative claims pattern, the beneficiary 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 beneficiary to meet the claims criteria for a given year and not the next year. See examples:

- Request A is submitted for 2014 claims for a cohort of beneficiaries with chronic kidney disease. The cohort is identified by applying the chronic kidney disease criteria to the universe of applicable claims for service provided on or before December 31, 2014 back through January 1, 2013 (a two-year reference period). A beneficiary meets the cohort inclusion criteria with one qualifying claim occurring in 2013 and has no subsequent claim meeting the specified criteria.
- Request B is submitted for 2015 claims for a chronic kidney disease cohort, which has a look-back period of December 31, 2015 back through January 1, 2014. Since the beneficiary's only qualifying claim occurred in 2013, the beneficiary does not meet the inclusion criteria. The beneficiary is included in the cohort for Request A but not Request B.

- *Modified Standard* definitions use the diagnosis and procedure codes defined in the standard request but allow researchers to modify other parameters such as beneficiary demographics, geography, and Medicare coverage status.

- *Custom* definitions allow researchers to request a cohort based on unique criteria provided by the researcher (e.g., all claims for a particular procedure, diagnosis, or specified population). This approach can also be used if the researcher used a different definition for one of the conditions already defined by the CCW classifications. This type of request may also include data requested based on a finder file using identifiers from a previous study.

The condition variables consider clinical criteria (from administrative claims), coverage criteria (from enrollment data), and specified time periods. The clinical criteria consider variations of the following:

- ICD-9/ICD-10, CPT4, or 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 Medicare Part A, B, and no HMO coverage. The specified time periods, or *reference time periods*, consider the length of time during which the clinical and coverage criteria are considered.

Three types of chronic condition variables are included in the Conditions segment and Other Chronic or Potentially Disabling Conditions segment of the MBSF. They include:

1. **Yearly**
   Algorithm criteria applied, using December 31 as the end of the reference year (e.g., 2015 yearly variable for algorithm with one-year reference period includes services between 01/01/15-12/31/15). The following are valid values for the yearly variable:
   0 = Neither claims nor coverage met
   1 = Claims met, coverage not met
   2 = Claims not met, coverage met
   3 = Claims and coverage met

2. **Mid-year**
   Algorithm criteria applied, using July 1 as the end of the reference year (e.g., 2015 mid-year variable for algorithm with 1-year reference period includes services between 07/01/14-06/30/15). The mid-year version of the conditions is only available for the 27 CCW conditions (not for the Other Chronic or Potentially Disabling Conditions segment). Researchers can use the MBSF to determine whether the beneficiary was alive and enrolled on July 1, for the purposes of producing statistics. The following are valid values for the mid-year variable:
   0 = Neither claims nor coverage met

1 = Claims met, coverage not met
2 = Claims not met, coverage met
3 = Claims and coverage met

3. ***Ever* date** (first occurrence of condition - YYYYMMDD)
   Date the beneficiary first met the *clinical* criteria of the algorithm (no coverage criteria applied), with the earliest possible date of 01/01/99.

   Values are null (missing) if the person never had a pattern of claims that indicate treatment for the condition. The earliest possible ever date for anyone in the CCW is January 1, 1999, although some beneficiaries obviously became eligible for Medicare before then. For beneficiaries who joined Medicare after that date, their ever dates will not precede the start of their Medicare coverage (i.e., the **COVSTART** variable in the MBSF).

Consideration of claims criteria for the algorithms includes a seven-day grace period for claim through dates occurring within:

- 7 days prior to first date of coverage
- 7 days after date of death

NOTE: Unless otherwise specified by the researcher, standard data requests will include (by default) all beneficiaries with yearly (or mid-year, if requested) variable = 1 or 3.

Control Populations

Requests for control populations should be made at the time of the initial data request. The inclusion/exclusion criteria for the control population should be specified by the researcher completing the data request form. Control populations may consist of those with particular conditions; alternatively, the researcher can request a control population lacking in any chronic conditions, if desired. Researchers can request a 1% or 5% sample file or customize the control population as needed. Specifications should include type(s) of data files, applicable diagnosis or procedure codes or DRGs, time periods, and any related demographic selection criteria.

Chapter 4. Medicare Claims Data Available through the CCW

The CCW includes Medicare enrollment and eligibility information, Medicare Institutional and Non-institutional claims, Medicare Part D enrollment and prescription drug fill events. The claims data files delivered from the CCW contain a subset of the original source files (see data dictionaries included in documentation information). Variables used infrequently or not applicable to a particular setting have been removed. A description of the variables contained within each file is located on the Data Dictionaries Page of the CCW website (http://www.ccwdata.org/web/guest/data-dictionaries).

A. Medicare Part A and B Claims

The Medicare claims found in the CCW are generally fee-for-service (FFS) Part A and B claims only (i.e., encounter information for services provided by MA plans is currently available only for 2015; Encounter records are contained a separate CCW data product). However, there are a few situations where the claims data does include services for Medicare Advantage (MA) enrollees. The two most notable instances are hospice care, which MA plans do not cover (it is always paid through FFS Medicare), and inpatient and skilled nursing facility services for beneficiaries enrolled in certain MA plans that are reimbursed based on costs and have the option of getting CMS to process those claims. The nuances of when services for managed care enrollees appear in Medicare's claims data are explained in greater detail in a technical publication from ResDAC (http://www.resdac.org/resconnect/articles/114).

- Structure of Claims

Institutional providers (such as hospitals, skilled nursing facilities, clinics, home health agencies, hospices, and outpatient dialysis facilities) bill for services using the 837I standard electronic format (previously known as the UB-04 claim form, or more recently as Form CMS-1450). Note that institutional providers use this form to bill for all services that they provide, regardless of whether the service is covered by Part A or Part B. Each claim has "base" and "revenue center" records. More information regarding the contents and processing of these claim forms is available on the CMS website; please see:

- http://www.cms.gov/Medicare/Billing/ElectronicBillingEDITrans/15_1450.htmlhttp://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/837I-FormCMS-1450-ICN006926.pdf

Non-institutional providers (such as physicians, other health care practitioners, and durable medical equipment [DME] providers) bill for services using the 837P electronic claim form (also known as the CMS-1500 claim form). These services are covered by the Part B benefit, and consist largely of professional services and DME. Similar to institutional claims, each non-institutional claim has "base" and "line item" records. More information regarding the contents and processing of these claim forms is available on the CMS website; please see:

http://www.cms.gov/Medicare/Billing/ElectronicBillingEDITrans/16_1500.html
Both kinds of claims are submitted to, and processed by, CMS contractors known as Medicare Administrative Contractors (MACs). As a historical note, the MACs have replaced separate entities that once processed institutional claims (known as Fiscal Intermediaries or FIs) and non-institutional claims (known as Carriers).

Note that Part D is not the only part of Medicare that covers outpatient prescription drugs; Part B covers certain drugs that generally must be administered in a medical setting, such as chemotherapy and some vaccines. Records for those drugs will appear as Part B claims rather than Part D events.

For both institutional and non-institutional claims, the ***base record*** contains the base or core portion of the claim. Each claim will also have at least one associated detailed record with more information regarding the particular services rendered. For institutional claims, those detailed records are known as ***revenue center records***; for non-institutional claims, they're known as ***line item records***. Both types of claims data have their own structure; see the CCW's website for source file record layouts and definitions (https://www.ccwdata.org/web/guest/data-dictionaries).

Additional details regarding use of the base claim or detailed revenue or line records is available in a CCW Technical Guidance document called "CCW Technical Guidance: Getting Started with CMS Medicare Administrative Research Files."

Table 16. Available CCW Medicare Data Files (Enrollment/Claims)

Files	Years	Linking Key or Stand Alone File
Medicare Enrollment Files		
Master Beneficiary Summary File (MBSF with CME as source)	2006 - current	CCW BENE_ID
Master Beneficiary Summary File (MBSF with EDB as source)	1999 – 2015	CCW BENE_ID
Institutional Claims Files		
Inpatient (IP) Base Claim Files	1999 – current	CCW BENE_ID
Inpatient Revenue Center Files	1999 – current	CCW CLM_ID*
Outpatient (OP) Base Claim Files	1999 – current	CCW BENE_ID
Outpatient Revenue Center Files	1999 – current	CCW CLM_ID*
Skilled Nursing Facility (SNF) Base Claim Files	1999 – current	CCW BENE_ID
Skilled Nursing Facility Revenue Center Files	1999 – current	CCW CLM_ID*
Home Health Agency (HHA) Base Claim Files	1999 – current	CCW BENE_ID
Home Health Agency Revenue Center Files	1999 – current	CCW CLM_ID*
Hospice (HOS) Base Claim Files	1999 – current	CCW BENE_ID
Hospice Revenue Center Files	1999 – current	CCW CLM_ID*
Reference Code Files - Delivered with Institutional Claim Files (note: separate file for each Institutional Setting – IP, OP, SNF, HHA and HOS)		

Files	Years	Linking Key or Stand Alone File
Institutional Condition Code Files	1999 – current	CCW BENE_ID
Institutional Occurrence Code Files	1999 – current	CCW BENE_ID
Institutional Span Code Files	1999 – current	CCW BENE_ID
Institutional Value Code Files	1999 – current	CCW BENE_ID
Demonstration/Innovation Code Files	2010 – current	CCW BENE_ID
Non-Institutional Claims Files		
Carrier Claim Files (Physician/Supplier)	1999 – current	CCW BENE_ID
Carrier Line Files (Physician/Supplier)	1999 – current	CCW BENE_ID*
Demonstration/Innovation Code Files	2010 – current	CCW BENE_ID
Durable Medical Equipment (DMERC) Claim Files	1999 – current	CCW BENE_ID
Durable Medical Equipment (DMERC) Line Files	1999 – current	CCW BENE_ID*
Demonstration/Innovation Code Files	2010 – current	CCW BENE_ID
Other CCW Medicare Files		
Medicare Part D Event Data**	2006 – current	CCW BENE_ID
Other Research Identifiable Files (CCW BENE_ID to HIC/SSN/Other crosswalk)	current	CCW BENE_ID

* The CCW_CLM_ID is the unique key to link revenue center information (for Institutional claims) or line item information (for Non-Institutional claims) to a specific claim.
** Medicare Part D Events and Part D Characteristics files are explained in detail in a separate user manual which is available on the CCW website.

- Final Action Status of Claims

Health care providers often submit more than one version of a claim for a particular service because they need to revise the information on the initial claim for some reason. For example, a hospital might need to revise the dates of service or diagnosis codes on an inpatient claim, or a physician might need to submit additional modifiers to specify the type of surgery that he or she performed. Any revision or adjustment requires a new claim (technically, most changes require two new claims, because the provider must submit one claim to cancel the initial claim and then submit another claim with the updated information), and some claims may be revised more than once.

The final action claim is the version of the claim where all adjustments to earlier claims have been resolved and CMS's final action on the claim is accurately recorded. Since weeks or months can pass between the provision of a service and the submission of the final action claim, CCW generally waits for the final action claim to appear before extracting data files for delivery to researchers. This period of time is often called the *run-out period*.

Claims are not considered final or complete until one year after the claim through date (CLM_THRU_DT). A full 12 months is allowed for claims to "mature" and the data files are considered final (i.e., a 12-month run out period). For 2005 forward, claims processed more than 365 days after the date of service are flagged in the data warehouse as *late arriving* records, and are not delivered. Over 99% of claims are processed within nine months of

service (note: this processing time varies slightly by care setting). For additional information, a CCW Technical Guidance paper which describes the level of claims maturity by processing month and setting is available on the CCW website (http://www.ccwdata.org/web/guest/technical-guidance-documentation).

To be consistent with the National Claims History (NCH), in May 2017 the CCW reprocessed the 2005 – 2017 Medicare Institutional claims using the revised CMS National Claims History final action (FA) algorithm.

The updated FA logic does not have a significant impact on overall claim counts or Medicare payment amounts. Total institutional claim counts and Medicare payments for 2015 changed by less than 0.03%. The largest impact was seen in inpatient psychiatric facilities, long term care hospitals, home health agencies, and other hospitals (e.g., cancer centers and children's hospitals). **Note that there is no impact from the FA update on carrier, durable medical equipment, Part D or enrollment data.**

- Claims Coding Systems

Health care claims vary in the kinds of information they require providers to submit about a patient's diagnoses and/or procedures obtained during a health care encounter.

Table 17 summarizes the types of information available in the various types of claims files.

Table 17. Diagnosis, Procedure, and Service Codes Used on Medicare Claims

Type of Code	Part A	Part B Institutional	Part B Non-inst.
ICD-9/ICD-10-CM diagnosis code	X	X	X
Diagnosis-related group (DRG)	X		
Revenue center code	X	X	
ICD-9/ICD-10-PCS procedure code	X		
Current Procedural Terminology (CPT) code	X	X	
Healthcare Common Procedure Coding System (HCPCS) code		X	X
Berenson-Eggers type of service (BETOS) code			X
Ambulatory payment classification (APC) code		X	

Medicare claims used the ninth version of the International Classification of Diseases (ICD-9) to classify all diagnoses, which identify the condition(s) for which a patient is receiving care, until September 2015. CMS switched to the next version of diagnosis codes, ICD-10-CM, for discharges and services on October 1, 2015. Claims data generally allow providers to specify numerous diagnosis codes (up to 25 codes for Part A claims and up to 12 codes for Part B claims – since 2011 when version "J" of the claim record was used; previously there

were up to 10 diagnosis codes for Part A and 8 for Part B), with one diagnosis identified on the claim as the principal or primary diagnosis.

Medicare pays for inpatient hospital care using case-mix groups known as diagnosis-related groups (DRGs), a classification system that groups similar clinical conditions and procedures. The beneficiary's principal diagnosis and secondary diagnoses, as well as any procedures furnished during the stay, are used to determine the appropriate DRG. CMS reviews the DRG definitions annually. The agency switched to a modified system, called Medicare Severity Diagnosis Related Groups (MS-DRGs) on October 1, 2007. Any claims that CMS received on or after that date are classified using MS-DRGs. Both DRGs and MS-DRGs appear in the same data field (CLM_DRG_CD) in the claims. The ICD-10 diagnosis codes have been mapped by CMS to the appropriate MS-DRG.

Medicare uses other forms of case-mix groups to pay for skilled nursing facility care (resource utilization groups, or RUGs) and home health (home health resource groups, or HHRGs). The RUG for SNF claims appear in the HCPCS_CD field (when the REV_CNTR code is 0022, then the first 3 digits of the HCPC_CD are the RUG). The HHRG for a particular revenue center (when the REV_CNTR code is 0023), is located in the data field called the ambulatory payment classification (APC) or Health Insurance Prospective Payment System (HIPPS) code (REV_CNTR_APC_HIPPS_CD). For more information about APCs, see: http://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/downloads/HospitalOutpaysysfctsht.pdf

Revenue centers are distinct cost centers within an institutional provider that can each submit separate charges. For example, most hospitals have distinct revenue centers for the emergency department, intensive care unit, physical therapy, laboratory, pharmacy, imaging, and so on. Revenue center codes (variable called REV_CNTR in the Revenue Center File for all Part A claims) are helpful for identifying different areas of the facility where the patient received care – and also other types of care which may affect payment (such as blood transfusions or laboratory tests).

For Part A inpatient hospital claims, providers used ICD-9 procedure codes to describe the specific procedures that they performed. Starting October 2015, the ICD-10 procedure coding system (ICD-10-PCS) is used in place of ICD-9 for procedure coding. For Part A claims that do not involve inpatient care and for Part B claims, providers use CMS Healthcare Common Procedure Coding System (HCPCS) codes to describe the services rendered (variable called HCPCS_CD). There are two levels of HCPCS codes. The first level are codes from version 4 of the Current Procedural Terminology (CPT-4), which is a numeric coding system maintained by the American Medical Association (AMA). The CPT consists of descriptive terms and identifying codes that are used primarily to identify medical services and procedures furnished by physicians and other health care professionals. The AMA makes decisions about adding, deleting, or revising CPT codes. The second level consists of

codes for procedures that are not included in the CPT codes; these are primarily used for non-physician services, such as ambulance services or durable medical equipment.

The Berenson-Eggers type of service (BETOS) classification scheme maps HCPCS codes into seven major categories (physician evaluation and management, procedures, imaging, tests, durable medical equipment, other services, and exceptions/unclassified services), with additional sub-categories within each. Additional details regarding use of BETOS codes (variable called BETOS_CD) in analyses is available in "CCW Technical Guidance: Getting Started with CMS Administrative Research Files."

B. Medicare Part D Prescription Drug Events

The CCW contains all Part D prescription drug events (PDEs), regardless of whether the beneficiary was enrolled in a managed care plan that includes coverage for prescription drugs or a stand-alone prescription drug plan. A detailed description of the Part D data is available in the "CCW. Part D Data User Guide", posted on the ccwdata.org website.

Chapter 5. Medicare Plan Characteristics

Starting with the 2015 benefit year, the CCW has prepared a Plan Characteristics set of six files per year with detailed information on those plan characteristics, which can be joined to the Medicare Beneficiary Summary File (MBSF) to better understand the benefits available to enrollees. Note that these Plan Characteristics files are an enhancement to the original CCW Part D Plan Characteristics files, which are available the for benefit years 2006 – 2014. The Plan Base, Premium, Plan Crosswalk and Service Area Files will now include Part C Only plans (i.e., plans that do not offer Part D prescription drug coverage – which are referred to as MA-PD plans) along with an indicator that allows users to subset by Medicare program type. In addition, a new file has been added that identifies the conditions that apply to each Special Needs Plan (SNP). The SNP Contracts file and the current Cost Sharing Tier file only include plans offering a Part D benefit.

The six files in the Plan Characteristics are:

1) The Plan Benefit Base File or "base" plan file contains key information about the managed care and/or drug benefit offered by the plan sponsor. Many of the variables in this file apply only to the Part D benefit and will be blank for Part C Only plans.
2) The Plan Premium File - has information on the premiums that each plan charged its enrollees. Most of the variables in this file only apply to plans that offer a Part D benefit.
3) The Plan Part D Cost Sharing Tier File - describes the features of the Part D plan benefit package, such as the tiers of the formulary, and has detailed information on how the cost of drug products will vary by benefit phase, the quantity of the drug dispensed, and the type of pharmacy used (e.g., in- or out-of-network).
4) The Plan Service Area File provides the regions included in the plan service area and has at least one row for every distinct plan ID and segment ID within a contract.
5) The Plan Crosswalk File will be useful to analysts interested in examining changes over time to the plans that were offered. Investigators can identify plans that are new, terminated in the prior year, were renewed or were consolidated with other plans. The file includes information for all plans that appear in the Plan Characteristics Files for the current year or the prior year.
6) NEW Special Needs Plans (SNP) Contracts file contains indicators to show which condition categories (e.g. heart failure, Diabetes] are covered in the SNP. SNPs are always MA-PD plans.

Chapter 6. Other CMS Data Available through the CCW

The CCW contains various types of CMS data from multiple care settings. In addition to Medicare enrollment and fee-for-service claims data files, Medicare Encounter data (for 2015), assessment data (*Minimum Data Set* and *Outcome and Assessment Information Set*), and Medicaid eligibility and claims data (delivered as MAX files) are available. A variety of CCW data files are displayed in Table 16.

For each Medicare beneficiary in the data file, the unique CCW identifier provides a common link across all applicable types of data available (variable called the CCW BENE_ID). Based on the approved research request, the CCW data delivered may or may not include patient identifying information. Regardless of whether patient identifying information is included, the unique beneficiary identifier provides researchers with the ability to analyze information across the continuum of care for a particular beneficiary or chronic condition cohort.

The unique CCW BENE_ID is created from the unloaded Enrollment Database (EDB) file, using the EDB Link Number, HIC number, and other beneficiary identifiers (i.e., gender, social security number [SSN]) for each unloaded EDB beneficiary. Analysis is performed to ensure that the beneficiary is not represented multiple times in the CCW HIC history table. If the beneficiary already exists in the CCW HIC history table, the table is updated with the new HIC information corresponding to the existing CCW BENE_ID. If they do not already exist in the CCW HIC history table, the record will be added and the beneficiary will be assigned a new unique CCW BENE_ID. All cross reference records for this beneficiary will be assigned this unique ID.

A. Medicare Encounter Data

In 2018, CMS announced the availability of 2015 Medicare Advantage (Part C) Encounter Data Research Identifiable Files (RIFs). These files are created from data submitted to CMS by Medicare Advantage Organizations (MAOs) that provide services to beneficiaries under the Medicare Part C benefit. Data dictionaries are on the CCW website (see https://www.ccwdata.org/web/guest/data-dictionaries).

B. Assessment Data

The Assessment and Medicaid data files available in CCW are identified in Table 18. Data dictionaries for all Assessment data files are on the CCW website (see https://www.ccwdata.org/web/guest/data-dictionaries).

C. Medicaid Data Files

- MAX Files

The Medicaid Analytic eXtract (MAX) data dictionaries, as well as details regarding the construction of the MAX files is available on the CMS website (https://www.cms.gov/MedicaidDataSourcesGenInfo/07_MAXGeneralInformation.asp). 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 MAX General Information on Data section of the CMS website. Information regarding some state-specific information is also available, such as managed care penetration. MAX file layouts and other helpful information regarding use of MAX files can be referenced on the ResDAC web site at (http://www.resdac.org/Medicaid/data_available.asp). A separate CCW technical guidance document provides many details regarding use of the MAX data. Please see "Getting Started with MAX Data Files – A Technical Guidance Paper" on the CCW website.

The MAX files are considered final once data for all states are ready, which is usually not until three or four years after the service/reference year.

Currently the MAX files for all states for 1999 – 2012, 28 states in 2013 and 17 states in 2014 are available. For the more recent years of MAX data fewer states have data, because starting in 2013 states began to transition from using the Medical Statistical Information System (MSIS) to the Transformed Medicaid Statistical Information System (TMSIS). As a result, many states are choosing to submit data to CMS using only the TMSIS format. Note that this transition to TMSIS is a major change for states that can take many months (or years) to fully implement.

- MMLEADS Files

In 2013, CCW began to offer a new data product designed for studying the Medicare and Medicaid dually enrolled population. The data files are called the Medicare-Medicaid Linked Enrollee Analytic Data Source (MMLEADS). They are a suite of analytical files consisting of a combination of linkable beneficiary, service and clinical condition data files. The MMLEADS beneficiary-level and service-level data files are person-level files. These files are intended to be linked together, as needed, for a variety of analytic purposes. Files are created as annual files from 2006 – 2012.

The MMLEADS files were produced from manipulated CCW Medicare and MAX source data and include enrollment/eligibility and summary claim/utilization and payment information. A data dictionary is available on the CCW website.

Table 18. Available Encounter Assessment and Medicaid Data Files

Files	**Years**	**Linking Key or Stand Alone File**
Medicare Encounter Files		
Inpatient (IP) Base Claim Files	2015	CCW BENE_ID
Inpatient Revenue Center Files	2015	ENC_JOIN_KEY*
Skilled Nursing Facility (SNF) Base Encounter Files	2015	CCW BENE_ID
Skilled Nursing Facility Revenue Center Files	2015	ENC_JOIN_KEY*
Home Health Agency (HHA) Base Encounter Files	2015	CCW BENE_ID
Home Health Agency Revenue Center Files	2015	ENC_JOIN_KEY*
Outpatient (OP) Base Encounter Files	2015	CCW BENE_ID
Outpatient Revenue Center Files	2015	ENC_JOIN_KEY*

Files	Years	Linking Key or Stand Alone File
Reference Code Files - Delivered with Institutional Encounter Files (note: separate file for each Institutional Setting – IP, SNF, HHA, and OP)		
Institutional Condition Code Files	2015	ENC_JOIN_KEY*
Institutional Occurrence Code Files	2015	ENC_JOIN_KEY*
Institutional Span Code Files	2015	ENC_JOIN_KEY*
Institutional Value Code Files	2015	ENC_JOIN_KEY*
Non-Institutional Encounter Files		
Carrier Encounter Files (Physician/Supplier)	2015	CCW BENE_ID
Carrier Line Files (Physician/Supplier)	2015	ENC_JOIN_KEY*
Durable Medical Equipment (DME) Encounter Files	2015	CCW BENE_ID
Durable Medical Equipment (DME) Line Files	2015	ENC_JOIN_KEY*
Assessment		
Minimum Data Set (MDS)	1999 – current	CCW BENE_ID
Outcome and Assessment Information Set (OASIS)	2000 – current	CCW BENE_ID
Swing Bed	2002 – current	CCW BENE_ID
Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI)	2002 – current	CCW BENE_ID
Medicaid Analytic eXtract (MAX) Files**		
PS – Person Summary File	1999 – current	CCW BENE_ID†
MESF – Medicaid Enrollee Supplemental File	1999 – current	CCW BENE_ID
IP – Hospital Inpatient File	1999 – current	CCW BENE_ID
LT – Long Term Care File	1999 – current	CCW BENE_ID
RX – Prescription Drug File	1999 – current	CCW BENE_ID
OT – Other Services File	1999 – current	CCW BENE_ID
Other Medicaid Analytic Files		
Medicare-Medicaid Linked Enrollee Analytic Data Source (MMLEADS)	2006 – 2012	CCW BENE_ID†

* The CCW- assigned ENC_JOIN_KEY is the unique key to link revenue center information (for Institutional encounters) or line item information (for Non-Institutional encounters) to a specific encounter record.

** MAX data are available for all states through 2012, 28 states in 2013, and for fewer than half of the states in 2014.

†Investigators may use the BENE_ID and STATE_CD or the Encrypted MSIS_ID and STATE_CD as the unique key for linking the MAX files.

Chapter 7. Format, Content and Encryption of CCW Output Files

This section describes the content and format of the CCW output package (the CCW data that are delivered to researchers). All User Guides can be referenced on the Analytic Guidance Page of the CCW website (http://www.ccwdata.org/web/guest/analytic-guidance).

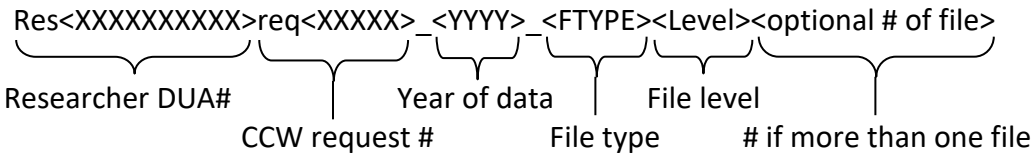
A. Format

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 0000077777, your CCW request number was 12345 and you obtained 2014 Part B carrier data, your folders and data files would look like this:

12345
2014
READ_ME_FIRST_REQ12345_2014.txt
res0000077777req12345_2014_BCARRB
res0000077777req12345_2014_BCARRL

The types of claims (or other data files) are identified using abbreivations.

- Medicare Beneficiary Summary File
  - Part A, B, C and D Segment – MBSF_ABCD_SUMMARY
  - Condition Segment – MBSF_CC_SUMMARY
- Assessments
  - MDS
  - OASIS
  - IRF

- Claims (see Table 19)

Table 19. Claim File Names

Claim Type	Claim/Revenue/ Line Files	File Level	Reference Code Files*	Demonstration Code File†
Inpatient	IP	B (base) or R (revenue)	INPT	INPT_DEMO
Skilled nursing facility	SNF	B R	SNF	SNF_DEMO
Hospice	HOSPC	B R	HSPC	HSPC_DEMO
Home Health administration	HHA	B R	HHA	HHA_DEMO
Hospital Outpatient	OP	B R	OTPT	OTPT_DEMO
Part B Carrier	BCARR	B (base) L (line item)	n/a	BCAR_DEMO
Durable Medical Equipment	DME	B L	n/a	DME_DEMO
Part D Event Data	PDES	n/a	n/a	n/a

* The reference code files include: _COND (Condition Code File), _OCCR (Occurrence Code File), _SPAN (Span Code File), and _VAL (Value Code File).

† The Demonstration code file is never populated prior to 2010.

The “B” at the end of the Institutional claim file names indicates that it is a base claim file; whereas the “R” identifies the corresponding revenue center. Similarly, for the Non-institutional claim files, the “B” at the end of the file name identifies the base claim file and file and the “L” indicates a line-item level file.

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 (e.g., res0000077777req12345_2014_BCARRL001).

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_REQ12345_2014.txt	
Claims Files		
File Name	**File Description**	**Unit of Analysis**
res<DUA number>_req<XXX>_<YYYY>_ IPB	Inpatient Base Claim File	Claim

res<DUA number>_req<XXX>_<YYYY>_ IPR	Inpatient Revenue Center File	Revenue Center Detail
Reference Code Files*		
File Name	**File Description**	**Unit of Analysis**
res<DUA number>_req<XXX>_<YYYY>_ IP_COND	Institutional Condition Code File (our example is for an Inpatient File)	Code Detail
res<DUA number>_req<XXX>_<YYYY>_ IP _OCCR	Institutional Occurrence Code File	Code Detail
res<DUA number>_req<XXX>_<YYYY>_ IP_SPAN	Institutional Span Code File	Code Detail
res<DUA number>_req<XXX>_<YYYY>_ IP_VAL	Institutional Value Code File	Code Detail
res<DUA number>_req<XXX>_<YYYY>_ INPT_DEMO	Inpatient Demonstration Code File	Demonstration Number Detail

* The Reference Code files (a set of 5 files) are routinely delivered for each of the institutional claim types (IP, SNF, Hospice, HH and OP). The Demonstration Code File exists for Carrier and DME files, in addition to all of the institutional claims files (2010+).

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 6C, below.

After you extract the data files, you should compare your record count to the control counts that 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>_ IPB. 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>_v6.sas	Program to read data into a SAS version 6.x environment. For example, the file inpatient_base_claims_read_v6.sas reads the inpatient base claims data into a SAS version 6.x environment.
<file name._v8.sas	Program to read data into a SAS version 8.x environment.

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

C. Encryption Information

The encryption technique for files extracted from the CCW uses the 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 CD, DVD, or USB 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.

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.

The CCW beneficiary identifier field (BENE_ID) is specific to the CCW (not applicable to any other identification system or data source). All requested data are linked using this field. It is encrypted using a cipher prior to delivery of data files to researchers. The Claim ID (CLM_ID) and Assessment ID (ASMT_ID) are also encrypted using the same cipher since these identifiers are also unique to a beneficiary. The encrypted BENE_ID is designed to be used by the researcher to link the data and the encrypted CLM_ID and ASMT_ID are to be used to identify records from the same claim/assessment). 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 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 8. Limitations of the CCW Data

There are certain expected anomalies in working with large, national, administrative datasets. Minimal data cleansing has occurred during processing of CCW data. However, some of the known limitations of CMS or CCW data are described below.

A. CCW Medicare Claims Data

Since claims (or encounter records) for most services provided to Medicare beneficiaries in managed care do not reach the claim data files, the CCW Medicare claims should be viewed as providing utilization information primarily for the FFS population.

Data files are received from CMS and loaded to the CCW on a monthly basis.
Claims are not considered final or complete until one year after the claim thru date. A full 12 months is allowed for claims to "mature" and the data files are considered final, as explained earlier in section 4A. Since data files are requested based on calendar years, researchers should consider the claims maturity or "completeness" of claims processing when requesting CCW data.

B. Assessment Data

Updated assessment records are obtained from the Quality Improvement and Evaluation System (QIES), 12 months after the completion date, and loaded to the CCW. CMS estimates that over 99% of assessments are processed within nine months of service (this processing time may vary by assessment type). Assessments may be updated until one year after the assessment date.

Beneficiary matching logic is applied, and the Assessments are populated with the person's BENE_ID. The presence of the BENE_ID enables the assessment records to be linked to other CCW Medicare and/or MAX data, if the DUA allows for this.

C. Invalid Values

Some of the CCW data files may contain invalid values, or values not conforming to the valid values provided in the CCW supporting documentation. The CCW data files contain data as received and processed from the original CMS processing source. Invalid values are processed, stored, and delivered as they are received. No modifications or conversions are made to "correct" for invalid variable values.

One exception is the removal of spaces or decimals to the left of diagnosis or procedure codes. Any periods or blank spaces occurring to the left of the first valid numeric value within a diagnosis or procedure code field have been removed. Diagnosis and procedure codes are stored without periods. No edits have been applied to remove any blank spaces occurring within a diagnosis or procedure code (between valid numeric values).

Chapter 9. Further Assistance with CCW data

The Research Data Assistance Center (ResDAC) offers free assistance to researchers using Medicare and Medicaid 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 CMS 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.

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

Appendix A - List of Acronyms

AMI – Acute Myocardial Infarction

BENE_ID – The unique CCW beneficiary identifier

CCW – Chronic Conditions Data Warehouse

CMS – Centers for Medicare & Medicaid Services

BASF – the Beneficiary Annual Summary File

BETOS – Berenson-Eggers Type of Service

BIC – Beneficiary Identity Code

BSF – Beneficiary Summary File

CAN – Claim Account Number

CDC – Centers for Disease Control and Prevention

CME - CMS Common Medicare Environment Database

CPT4 – Current Procedural Terminology

DME – Durable Medical Equipment

DMERC – Durable Medical Equipment Regional Carrier

DOB – Date of Birth

DOD – Date of Death

DRG – Diagnosis Related Group

DUA – Data Use Agreement

ER – Emergency Room/Department setting

EDB – CMS Enrollment Database

ESRD – End Stage Renal Disease

FFS – fee-for-service claims

HHA – Home Health Agency

HIC – Medicare beneficiary Health Insurance Claim number

HMO – Health Maintenance Organization

HOS – Hospice

HOP – Hospital Outpatient

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

IRF – Inpatient Rehabilitation Facility

IP – Inpatient Hospital

IPF – Inpatient Psychiatric Facility

IRF/PAI -- Inpatient Rehabilitation Facility Patient Assessment Instrument

LIS – Low-income subsidy

LT – Long Term Care

LTC – Long Term Care Hospital

MA – Medicare Advantage

MCBS – Medicare Current Beneficiary Survey

MAX – Medicaid Analytic eXtract data files

MBSF – CCW Master Beneficiary Summary File

MCBS – Medicare Current Beneficiary Survey

MDS – Minimum Data Set

MESF – Medicaid Enrollee Supplemental File

MMLEADS – Medicare-Medicaid Linked Enrollee Analytic Data Source

NDI – National Death Index

OASIS – Outcome and Assessment Information Set

OP – Hospital or other Institutional Outpatient Setting

OT – Other Services

PDE – Part D prescription events

PGP – Pretty Good Privacy

PLRO – patient liability reduction due to other payer

PS – Person Summary

ResDAC – Research Data Assistance Center

RTI – Research Triangle Institute

RX – Prescription Drug

SDA – Self-Decrypting Archive

SNF – Skilled Nursing Facility

SSA – Social Security Administration

TrOOP – True Out-of-Pocket