Evaluating Encounter Data Completeness

For Researchers using the Centers for Medicare & Medicaid Services Chronic Condition Data Warehouse (CCW)

Date: May 9, 2012

Table of Contents

Abstract

Under traditional Medicaid fee-for-service (FFS) delivery systems, rendering providers overwhelmingly submit their bills directly to each state's Medicaid program in order to receive payment. Over the past two decades, Medicaid managed care delivery models have been implemented for specific populations, services or geographic areas in over half of all states. As Medicaid programs have implemented these managed care delivery models, the need for complete and accurate encounter data has become increasingly critical for states to understand and manage access, quality and cost within their Medicaid programs. Unlike FFS delivery systems, under risk-based managed care models the rendering providers do not submit their billing information directly to the state; instead providers submit their billing information to managed care organizations (MCO's); and after the claim is processed, the MCOs are required to submit their detailed claims to the state as encounters. The encounter data is then used by the state to manage access, quality and cost and to make policy and program design decisions.

The encounter data process is complicated and is susceptible to error. As a result, it is important to identify any potential gaps in the data before using it for program analyses or for general research purposes. There are multiple strategies available to measure encounter data completeness. Using these strategies, a researcher can determine whether the data is valid or needs to be researched further before using. In this report macro validation is used as a strategy to validate encounter data from the Chronic Condition Data Warehouse (CCW). Metrics were calculated from the data and used to measure completeness across states and Medicaid populations.

Keywords: Medicaid, Managed Care, Encounter Data.

Encounter Data Overview

In order to analyze population, cost, and utilization characteristics of Medicaid programs, researchers need to study both fee-for-service (FFS) claims data for services paid for directly by the states and encounter data for services paid for through a Medicaid managed care organization. In 2010, 36 state Medicaid programs used risk-based managed care arrangement to deliver services to over 26 million members.[1] In this delivery model, state Medicaid programs pay managed care organizations (MCOs) a monthly capitation rate per member (also known as per member per month or PMPM). This capitation rate covers a defined set of benefits that the MCOs are responsible to deliver, including medical services, pharmacy, long-term care, and other services based on their contract with the state. Rendering providers are paid for these services based on claims submitted to and processed by the MCOs. In order for a state to evaluate the managed care portion of their delivery system, state officials require the MCOs to provide detailed information on the use of health care services by their members, known as encounter data. Encounter data provides detailed health care information on the specific health care services utilized by the MCO's Medicaid beneficiaries. In some states, MCOs are required to submit encounter data that provides the same level of detail as FFS claims data.[2]

Studying encounter data is the most direct route toward understanding access and quality of care received by Medicaid members enrolled in risk-based managed care. For example, the diagnosis information contained can indicate the conditions being treated, which can lead to development of disease management or other programs to help beneficiaries better control their chronic conditions. In addition, many state Medicaid managed care programs determine the severity of its population based on diagnosis data, then risk adjust their MCO capitation rates to ensure a fair payment rate. An MCO with a higher-risk population will receive a higher capitation rate than an MCO with a lower-risk population. Encounter data can also be used to help measure the quality of care and underlying utilization patterns in a Medicaid program using measures such as the Healthcare Effectiveness Data and Information Set (HEDIS). States and MCOs can also use encounter data to profile providers within the Medicaid program, allowing a health plan to benchmark providers against their peers, in order to identify deficiencies and improve quality of care.

Encounter Data Completeness and Accuracy

Capturing, sending and receiving encounter data has historically been difficult and costly for MCOs and states alike. The encounter data process is lengthy and has many steps where data can be lost or errors can be introduced into submitted data elements. States and MCOs have developed a number of validation techniques to helping to ensure the completeness and accuracy of their encounter data. Some of these various encounter validation techniques are described below.

---

[1] "A Profile of Medicaid Managed Care Programs in 2010: Findings from a 50-State Survey", Health Management Associates; Kathleen Gifford, Vernon K. Smith, and Dyke Snipes and Kaiser Family Foundation; Julia Paradise; September 2011. http://www.kff.org/medicaid/upload/8220.pdf

[2] "Collecting Reporting and Using Encounter Data: A Primer for States", Mathematica Policy Research; Vivan L.H. Byrd and James Verdier; October 2011

THE LEWIN GROUP®
543483

GENERAL DYNAMICS
Information Technology

Measuring Encounter Data Completeness

Depending on the resources available and the structure of the managed care delivery system, there are multiple strategies that may be used to accurately measure encounter data completeness. Regardless of the strategy used, encounter data completeness should at a minimum be measured each year, as data submissions could vary from year to year. When a state first implements a managed care program they should evaluate their encounter data on a monthly basis to help identify deficiencies for their MCOs and work closely with MCOs to correct any observed deficiencies.

Macro Validation

Macro validation involves testing the entire "universe" of data to be studied. All of the encounter data is summarized and metrics are developed which can be compared both within and across states, across MCOs and over time. Macro validation can be performed through an automated process and can be repeated as new encounter data comes in. This should be the first step in the validation process. Once the overall validation is completed, more targeted studies can be performed (e.g., examining service delivery among specific providers or provider groups and/or for subpopulations of members, such as children with asthma or those with multiple visits to the emergency department within a given year).

One common macro validation strategy evaluates the number of unique recipients receiving services. The number of unique beneficiaries receiving a service can be counted for each category of service (e.g., physician, inpatient hospital pharmacy). These counts can be studied over time to determine if there are gaps in the study population (state, MCO, etc.) for specific service months. However, in order to make comparisons to other populations, the data should be normalized. To do this, the counts are divided by the number of eligible beneficiaries in the study population. The end result is a percentage of eligible beneficiaries that had an encounter.

Micro Validation

Micro validation involves evaluating the completeness for individual providers and/or beneficiaries. On a provider level, this includes developing primary care provider profiles and evaluating services used by members in a provider's panel. For example, a researcher can look at members with diabetes in each provider's panel and see whether they are receiving appropriate drugs, laboratory tests and check-ups. Although there are certainly reasons why individuals would receive more or less care, across a panel of patients the services should be relatively consistent. On a member level, this includes comparing individual encounters to medical records to determine whether the encounter data includes the correct diagnosis information and services provided. These analyses require additional effort to complete. They can be undertaken if anomalies are found through the macro validation analysis. This strategy is typically employed by states and other payers to validate the information reported by participating health plans to verify HEDIS measures and other plan reported data. This strategy is normally not used by researchers because of the resources required to conduct the validations.

Financial Validation

State Medicaid agencies generally require participating MCOs to submit periodic reports that contain both cost and utilization metrics for covered beneficiaries. When working with a specific state, it is possible to summarize the cost from the encounter data and compare it to the reported financial data. In order to do this, the financial data should represent raw payments to providers and should not include any adjustments, including completion factors to estimate current cost. The encounter data should represent the same service date timeframe used in the reports. The encounter data will then need to be summarized. The service categorization used when summarizing the encounter data should be consistent with that used to develop the reports to allow for like comparison. The financial reports will need to come from the state Medicaid agency, as they may not be publicly available. This strategy is normally performed by actuaries in the process of establishing MCO capitation rates.

Measuring Chronic Condition Warehouse Encounter Data Completeness

To evaluate Medicaid programs across the United States, researchers can use the Chronic Condition Data Warehouse (CCW) which includes Medicaid eligibility data as well as FFS data and managed care encounter data. The MCO submits its encounter data to the state's Medicaid Medical Information System (MMIS). The data is usually run through a series of edits which may deny encounters based on beneficiary eligibility or errors in submitted data elements. The remaining data is then submitted to CMS's Medicaid Statistical Information Statistics (MSIS) and formatted to create the annual Medicaid Analytic eXtract (MAX) data files. The CCW is populated with data from MAX. As a result of the potential for data issues either at rendering health care provider offices, MCO data processing systems, or state MMIS systems, it is especially important to determine the completeness of the CCW encounter data before using it for analysis. Additional data checks can be performed for each state by reviewing the MAX validation and anomalies reports that are published each year. These reports provide detailed findings of errors and anomalies that have been found in the MAX data for that year.

We used a macro validation strategy based on physician office visits when working with the 2008 CCW encounter data for this report. To complete this process, both encounter and eligibility data were extracted from the CCW and utilization metrics were calculated. Analysis of physician office visit rates yields insight into encounter completion rates due to the fact that within like populations, physician office visit rates should be relatively stable. Our completion strategy evaluated office visit rates for each state operating a full risk-managed program; we compared the rates for the same full risk-managed care populations across states. This is one example of a macro validation strategy that can be used to validate encounter data completeness. A similar comparison can be used within a population over time to identify periods of time when the data may be less complete due to systems issues within individual MCOs or state MMISs. There are numerous strategies that should be used to evaluate encounter data completeness prior to conducting any analysis of a state's managed care program. In addition to evaluating physician services, studies should also be conducted of inpatient, emergency room, pharmacy and other major services. In addition to the volume of encounters, studies should be conducted of the accuracy and completeness of specific data elements that will be used in your research such as diagnosis, procedure and National Drug Code (NDC) codes. Risk adjustment strategies can also be employed that compare the prevalence of chronic

conditions identified by diagnosis codes with the prevalence rates for chronic conditions identified using pharmacy data.

Extracting Encounter Data

For purposes of this study, claims and encounter data for services delivered in 2008 were extracted and loaded into state-specific datasets. The first step was to extract the detailed encounter data from the CCW. This included the Inpatient Record (IP), Drug Record (RX), Long-Term Care Record (LT) and Other Services Record (OT). The physician encounters are included in the OT record.

For our analysis, a hierarchy was used to assigned encounters to a specific service category. This was done using procedure codes reported on the encounters. Some encounters have multiple detail lines which are separate records in the CCW data but still part of the same encounter. Since the CCW data for 2008 does not contain a claim number, an encounter was defined by all records for the same member identification number, provider identification number and beginning date of service. To categorize all of the records reported on an encounter to the same service category, a hierarchical list of service categories (i.e., office visits, therapies and lab/x-ray) was developed. The office visit category was defined using a list of industry accepted procedure codes for these services. The physician office visit grouping is at the top of the hierarchy. If a physician office visit procedure code exists on any single record on an encounter, then the entire encounter is categorized as a physician office visit. Only physician office visit encounters were retained for further study.

Measuring Membership

In addition to the encounter data, developing meaningful and comparable metrics also requires the total number of members by month, which will serve as the denominator. Membership data is stored in the Personal Summary Record (PS) within CCW and includes a record for each unique person. This record includes demographic information (e.g., eligibility classification, gender, date of birth, and ethnicity). Several categories were derived from the PS data for use in the encounter data completeness analysis:

Eligibility Category: The eligibility group is determined by mapping the CCW Uniform Eligibility Code field to high-level groupings, e.g., Temporary Assistance to Needy Families (TANF) and Supplemental Security Income (SSI). Medicare codes are used to determine Medicare and Medicaid dual eligibility. Detailed logic describing the use of eligibility codes on the CCW PS file that were used to create the eligibility groups is provided in Appendix A.

Age Category: The age category is determined by mapping the Eligible Age Group Code field to one of the following groupings: Newborn, Child, Adult, Elderly, and Unknown.

MCO: This indicates if a member was enrolled in a full-risk managed care program. Managed care membership is counted based on the Eligible Pre-Paid Plan Months Count. For each specific member, the Eligible Pre-Paid Plan Code is used to determine managed care eligibility in a given month. If a member was not enrolled in a managed care organization during a specific month, then that member was not included in the study for

that month. Detailed logic describing the use of plan codes to assign a member's managed care status is provided in Appendix A.

Calculating Metrics

Once the encounter and eligibility data was extracted, the resulting datasets were combined based on the MSIS Identification (MSIS ID) Number field, which identifies the specific eligible member, and the month of eligibility/month of service. Encounters that did not have an MSIS ID that matched to the eligibility file for the month of service were removed from the study.

Using the remaining encounter data, the following values were calculated by month, eligibility category, and age category:

- **Total Eligible Members** – Count of total eligible members based on MSIS ID;
- **Total Claims** – Count of office visits (unique member, provider and date of service); and
- **Total Users** – Count of unique members with an office visit (members with multiple visits are counted once per month).

From those values, the following metrics were calculated:

- **User Percentage** – The percentage of members that had an office visit during each month of service; and
- **Claims per User** – The average number of office visit encounters per unique user during each month of service.

As discussed above, user percentage will be the primary metric that will be used to make comparisons among states. States and populations with very low values, especially in comparison to other states, should be investigated in more detail.

## CCW Encounter Data Completeness Results

Every state with 2008 data in the CCW was extracted and studied. The eligibility and age categories with the highest membership (TANF and SSI) were studied further.

TANF

TANF is the largest eligibility category in Medicaid. Based on CCW data, there was an average of 7.5 million TANF members per month enrolled in full-risk managed care. This represents 58 percent of the total full-risk managed care membership. 6.2 million TANF members are children age 1 - 20. 1.2 million TANF members are adults age 21 - 65. The remaining 100 thousand members are either less than 1 or over 65. Figure 2 shows the office visit user percentage by state for the TANF adult and child populations.

Figure 2: Percent of Managed Care TANF Children and Adults with Physician Encounters in 2008
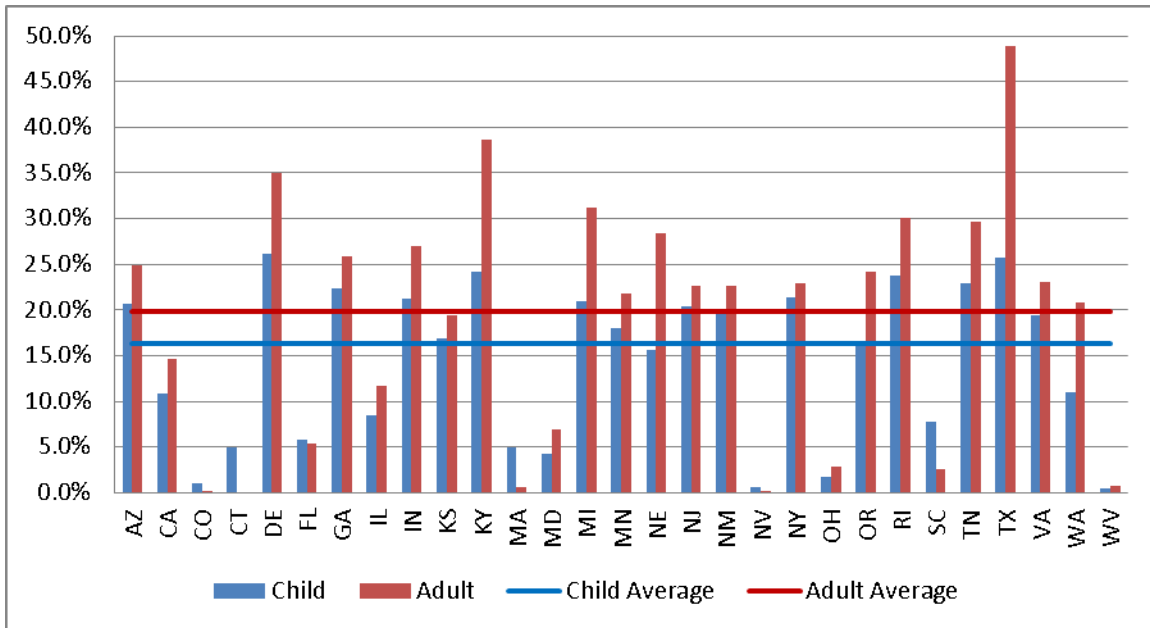
The average user percentage rate was 16.3 percent for children and 19.9 percent for adults. This difference is expected, as adults typically have higher medical service utilization than children.
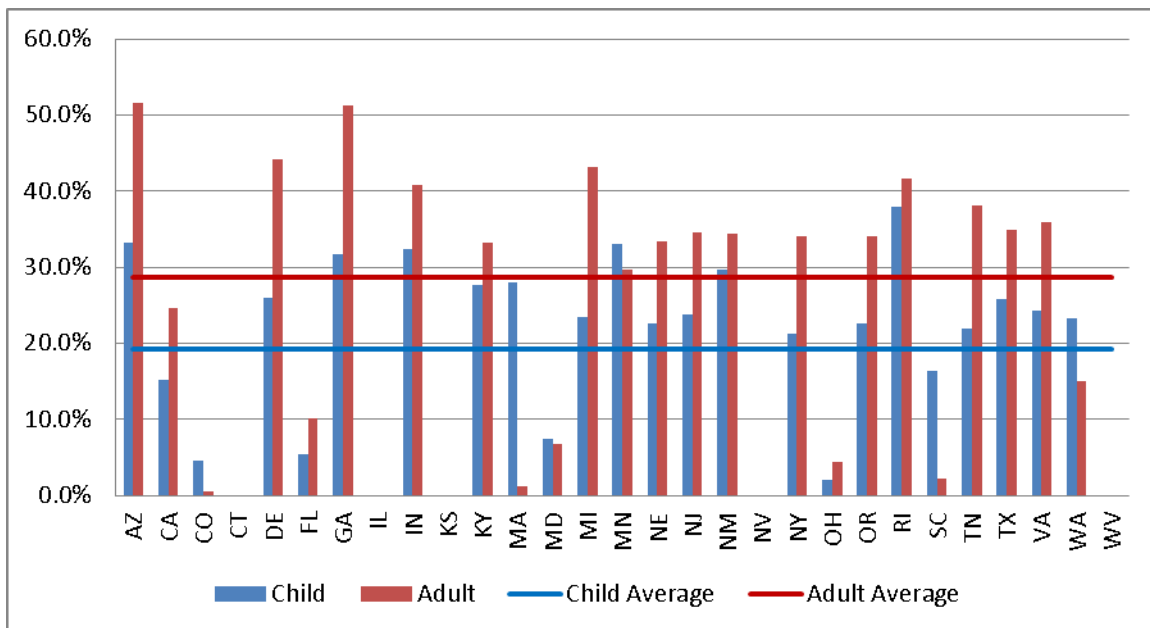
Observations

- The average national rates were partially driven by California, which has 1.4 million children and 466 thousand adults. Both California percentages, 11 percent and 15 percent respectively, had the effect of driving down the average.

- Some states with active full-risk managed care programs, such as Nevada, had virtually no encounter data. This could be due to a data issue occurring in the state's submission to MSIS.

- States such as Michigan, New York, and New Jersey have strong membership information for both child and adult populations. The encounter data for these states also looks complete in aggregate and should be sufficient to support further analyses of the managed care programs in these states. The completeness of the encounter data could be further investigated by examining user percentages for other service categories in addition to the physician office visit completion analysis discussed above, as well, completing similar user percentages analyses by region or for individual MCOs.

- States such as Massachusetts, Maryland, and South Carolina have large MCO enrollment but extremely low utilization. The encounter data for these states looks incomplete and is not sufficient to support further analyses of the managed care programs in these states. The completeness of the encounter data could be further investigated by examining user percentages for other service categories and examining user percentages by region, or for individual MCOs to determine if the data is complete for select plans or regions.

SSI

SSI is another large Medicaid eligibility category. The majority of SSI beneficiaries, however, have both Medicare and Medicaid coverage. Medicaid is the payer of last resort, and it is common to have no Medicaid physician office claims for these beneficiaries since office visits are covered under Medicare. As a result, this study focused on the Medicaid-only population which was identified using eligibility codes as referenced in Appendix A. Based on the CCW data, there was an average of 945 thousand SSI Medicaid-only members per month enrolled in full-risk managed care. This represents seven percent of the total full-risk managed care membership. Figure 3 shows the office visit user percentage by state for the SSI adult and child populations.

Figure 3: Percent of Medicaid-Only SSI Children and Adults with Physician Encounters in 2008


The average user percentage rate was 19.2 percent for children and 28.6 percent for adults. The number of visits is higher for SSI than TANF because SSI beneficiaries are more likely to be elderly, blind and/or have disabilities.

Observations

- States such as Michigan, New York, Texas and Virginia have significant Medicaid managed care membership and above average user percentage. These states can be investigated further for use in research.

- States such as Ohio, Maryland and Massachusetts should be researched further prior to use, as both SSI age categories have significant Medicaid managed care membership but extremely low user percentages. The encounter data for these states looks incomplete and is not sufficient to support further analyses of the managed care programs. The completeness of the encounter data could be further investigated by examining user percentages for other service categories and examining user percentages by region, or for individual MCOs to determine if the data is complete for select plans or regions.

The Lewin Group evaluated each state to determine the usability of its encounter data. Figure 4 ranks the states in four categories. Green states have physician utilization results above the average for multiple populations. Yellow states have results slightly below average or above average for one population and below average for another. Red states have results significantly below the average for multiple populations. White states either do not have managed care programs or do not have encounter data loaded in the CCW, and as a result, were not studied.

Figure 4: Encounter Data Usability by State
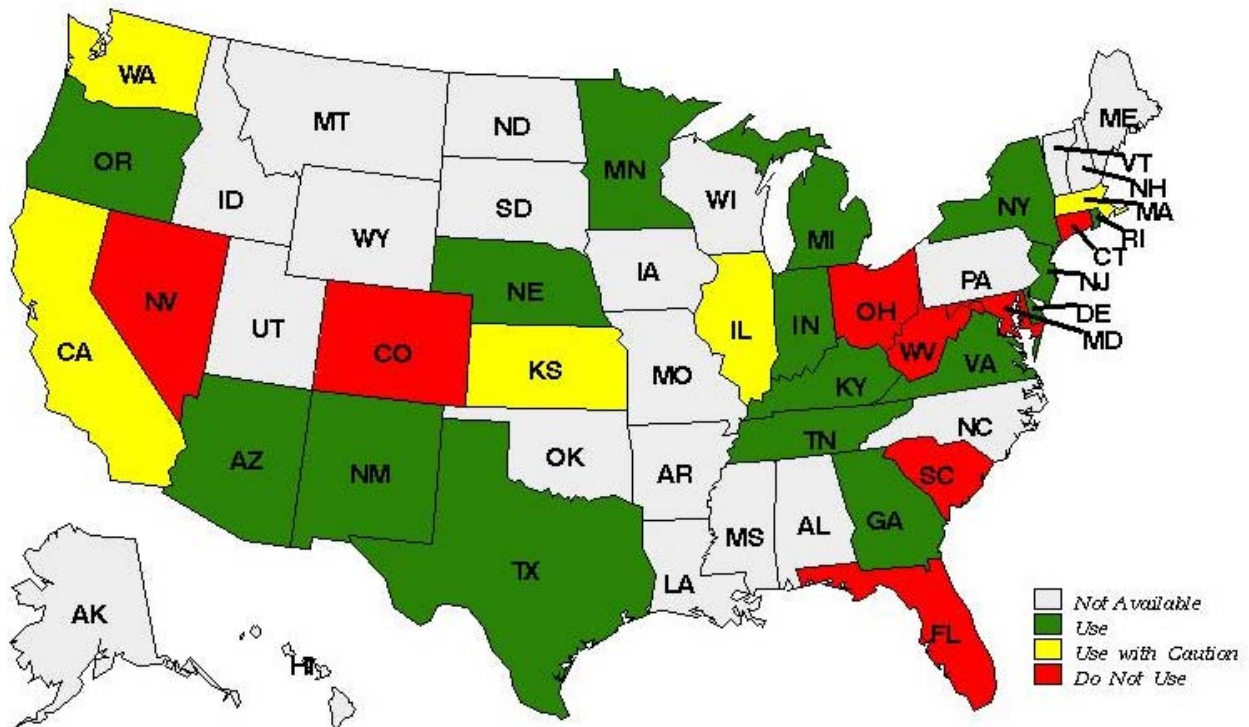
Conclusion

The CCW is a powerful one-stop source of detailed Medicaid data which researchers can use to gain insight into the Medicaid population. It provides information on demographics, utilization and cost of medical services. Like any data source, however, there can be data quality considerations that need to be identified and reflected on a case by case basis based on the state and research design. Medicaid encounter data provides additional opportunity for error based on the complicated nature of the process involved in submitting the encounter data from the rendering provider's office to the MCO and then to the state MMIS system. As a result, it is important to perform encounter data completeness validation studies before making any assumptions or conclusions from the data. Our analysis revealed multiple states with extremely low user percentages. These results are most likely indicative of encounter data completeness problems and not lack of access to physician services for Medicaid members in those states. In

this case, using the CCW encounter data to evaluate the performance of the managed care program and participating MCOs in those states will be misleading.

Appendix A
Derivation of a Member's Eligibility Category and Managed Care Status

The Chronic Condition Data Warehouse (CCW) Personal Summary (PS) file was used to determine a member's eligibility group, age category, managed care enrollment and their number of months of Medicaid eligibility during 2008. The CCW variables and the value for each variable used to assign members to categories are provided in the sections below.

Eligibility Category Assignment

The latest CCW eligibility code reported for a member (EL_CCW_ELGBLTY_CD_LTST) was used to assign the member to an eligibility group for those members that were not dually eligible for Medicaid and Medicare. For dually eligible members, the Medicare Dually Eligible code (EL_MDCR_DUAL_ANN) and their original Medicare reason for entitlement (MDCR_ORIG_REAS_CD) was used to assign them to an eligibility group. In the assignment process logic, dually eligibles were assigned first. If a member was not determined to be dually eligible they were assigned to another eligibility category.

The values that were used to identify dually eligible members are provided in Table 1. Several methodologies have been used by users of the CCW data to identify dually eligible recipients. This methodology was selected based upon feedback from other departments within CMS. Researchers may want to consider other methodologies and should discuss alternatives with their project teams. The values contained in the CCW may also change in future years.

Table 1. Variable/Values Used to Identify Medicare-Medicaid Dual Eligible Enrollees

Dual Eligible Category	EL_MDCR_DUAL_ANN values	MDRC_ORIG_REAS_CD values
Dual- Partially Eligible (i.e., QMB/ SLMB)	01, 03, 05, 06, 07, 51, 53, 55, 56, 57	Not used in assignment process
Full Dual - Aged	02, 04, 08, 52, 54, 58	0
Full Dual - Disabled	02, 04, 08, 52, 54, 58	1, 2, 3

The values that were used to identify non-dually eligible members to eligibility categories are provided in Table 2. There are numerous values for the eligibility variables reported in the CCW data that have been derived from state specific eligibility codes. This methodology was selected based upon feedback from other departments within CMS. Researchers may want to consider other methodologies and should discuss alternatives with their project teams. The values contained in the CCW will also be updated in future years.

Table 2. Variable/Values Used to Identify Non-Dual Eligibility Categories

Eligibility Category	EL_CCW_ELGBLTY_CD_LTST values	EL_CHIP_FLAG_latest *
SSI	11, 12, 41, 42	
TANF	14, 15, 16, 17, 34, 35	
MA- Only, SSI Related	21, 22	
MA – Only, Non SSI Related	24, 25, 44, 45	
Foster Care	48	
Waiver	51, 52, 54, 55	
CHIP	14, 15, 16, 17, 34, 35, 44, 45, 54, 55	2, 3

- The most recent monthly value for the EL_CHIP_FLAG series of variables was used to determine a member's CHIP status.

Age Category Assignment

A member's age group code (EL_AGE_GRP_CD) was used to assign them to four age categories; Newborn, Children, Adult and Elderly. The values that were used to assign members to each category are provided in Table 3.

Table 3: Variable/Values Used to Identify Age Category

Age Category	EL_AGE_GRP_CD values
Newborn	0
Child	1, 2, 3
Adult	4, 5
Elderly	6, 7, 8
Unknown	Any other value

Managed Care Status

Members enrolled in MCOs were identified using the MC_COMBO_MO_ series of variables. In a given month, if the MC-COMBO_MO variable had a value of 1, 6, 7, 8, or 9 then a member was classified as being enrolled in a comprehensive managed care organization.