Chronic Conditions Warehouse

Your source for national CMS Medicare and Medicaid research data

Chronic Conditions Warehouse Virtual Research Data Center

Medicare Risk Score Operational Payment Base and Model Output Files User Guide

JUNE 2026 | VERSION 1.1

Revision Log

Date	Changed By	Revisions	Version
June 2026	D. Happe	Updated ResDAC contact information to their contact webpage	1.1
April 2026	K. Schneider S. Pietzsch	Created initial document	1.0

Table of Contents

1.0 Overview 1

2.0 Background 2

3.0 CCW Payment Risk Score Files 4

3.1 Source Data 4

3.2 CMS HCC Data Files 5

4.0 CCW Payment Risk Score Files 6

4.1 CCW Operational Payment Risk Score Base Risk Score File 6

4.2 CCW Detailed MOFs 9

4.2.1 ESRD V21 Scores 9

4.2.2 V21 Scores 10

4.2.3 V22 Scores 10

4.2.4 V23 Scores 11

4.2.5 V24 Scores 11

4.2.6 RXHCC Scores 11

5.0 Medicare Risk Scores 12

5.1 Calculation of Risk Scores 12

5.2 Hierarchical Condition Categories 13

6.0 Medicare Enrollment Data 15

7.0 Format, Content, and Encryption of CCW Output Files 16

8.0 Further Assistance with CCW Data 18

Appendix A — Acronyms 19

List of Tables

Table 1. Base file — resolved payment risk scores 6

Table 2. Model segment code values, by payment year 7

Table 3. HCC model segments by payment year 9

Table 4. Naming conventions for HCCs — V21 example 10

Table 5. Variables added together to create the raw risk scores within the risk adjustment model 12

Table 6. Adjustments within the risk adjustment model 13

Table 7. Format and naming convention for the CCW physically delivered files 16

Table 8. Payment risk score files in the CCW VRDC 17

List of Figures

Figure 1. Risk adjustment processing to create research risk score files 4

1.0 Overview

Medicare is the primary health insurance program for people aged 65 or older, people under age 65 with disabilities, and people of all ages with end-stage renal disease (ESRD). Centers for Medicare & Medicaid Services (CMS) developed risk scores that Medicare uses to adjust payments to Medicare Advantage (MA) plans to account for differences in health status among plan enrollees.[1]

CMS uses the Chronic Conditions Warehouse (CCW) to develop and manage CMS research data resources. The CCW has complete (100 percent) Medicare enrollment and fee-for-service (FFS) claims data, obtained directly from CMS. CCW also obtains Medicare payment risk score data files from CMS. From this source data, the CCW team has prepared data files to disseminate to researchers 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 (refer to the Research Data Assistance Center (ResDAC) website for information on requesting Medicare data).

The Medicare Risk Score Operational Payment Base and Model Output files (MOFs) are available from the CCW starting with the 2019 payment year. This information is available for all Medicare beneficiaries, regardless of whether the beneficiary enrolled in Original Medicare or MA. CMS designs the files for investigators to use them with other CCW data products, such as the Master Beneficiary Summary File (MBSF) that contains a wide range of demographic and Medicare coverage information. When the CCW team loads the source data to the CCW database, the team assigns each record the CCW unique beneficiary identifier (the BENE_ID variable), to facilitate linkage between files and for longitudinal analyses.

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

[1] CMS. Medicare, Risk Adjustment webpage. (Accessed 03/06/2026)

2.0 Background

Differences in beneficiaries' health status can have a significant influence on their expected health care use and spending, or payment risk. One tool that can help understand those differences is the risk scores that Medicare uses to adjust payments to MA plans to account for differences in health status among plan enrollees. CMS expects MA plans with enrollees who have higher than average risk scores to incur higher than average costs; therefore, CMS adjusts the capitated payments to MA plans accordingly. Those risk scores are based in part on diagnosis codes that CMS then groups into related disease categories known as Hierarchical Condition Categories (HCCs).

The CMS-HCC risk adjustment model factors for Medicare Part C are calibrated on the Medicare OM population, and thus rely on expenditures and diagnoses from FFS claims to calculate the predicted costs associated with each of the risk adjustment model factors.[2] Although CMS has begun using MA plan-submitted diagnoses from encounter data to calculate risk scores, the Part C risk adjustment model coefficients continue to be calibrated on the OM population. CMS calibrates Part D risk adjustment models using prescription drug event (PDE) data from both OM and MA-PD enrollees.

The HCC-based payment risk scores are estimates of the projected costs for the beneficiary relative to the national average relative FFS costs. The purpose of risk adjustment is not to accurately predict costs for any particular person, but on average for a group of beneficiaries with the same attributes that affect health care costs.[3]

The average (normalized) risk score for the overall FFS Medicare population is set at 1.0. CMS expects beneficiaries with scores greater than 1.0 to have above-average Medicare costs (and proportionally higher as well i.e., a score of 2.0 means CMS expects costs to be twice the overall average), and vice versa.[4] Risk scores are based on a combination of factors: the beneficiary's age and sex; whether the beneficiary is dually enrolled in Medicaid, first qualified for Medicare on the basis of disability, or lives in an institution (usually a nursing home); and the diagnoses on the beneficiary's claims from the previous year if the beneficiary was enrolled in Part B for all 12 months of the previous year (i.e., the risk score for payment year 2019 is based on diagnoses from claims from 2018). Diagnosis codes are not a model input for beneficiaries with less than 12 months of Part B during the previous year. CMS employs a "new enrollee" model for these beneficiaries.

CMS pays each MA organization a monthly amount per beneficiary enrolled in its plan. Each year, CMS evaluates and updates the HCC model to incorporate changes in the diagnosis codes that map to HCCs.[5] Periodically, CMS makes significant structural changes to improve predictive accuracy of the HCC model for subpopulations. For example, in 2017, CMS implemented version 22 (V22) of the HCC model.[6] It includes additional segments to better address disease patterns and cost differences between the aged versus disabled subpopulations and by status of Medicare-Medicaid dual eligibility (non-dual, partial benefit dual, or full benefit dual). In 2019, CMS implemented the V23 model, with resolved risk score reflecting a blend of the V22 and V23

---

[2] In the 2015 payment year, CMS began to use risk adjustment eligible diagnoses from encounter data submissions in the calculation of risk scores used for MA risk adjusted capitation payments. CMS phased in the use of diagnoses from encounter data to calculate MA risk scores over a multi-year time period. Over this multi-year phase in, CMS transitioned the source of MA diagnoses used to calculate MA risk scores from those submitted to the legacy Risk Adjustment Processing System (RAPS) to solely using MA diagnoses submitted on encounter data (using the X12 837 5010 format that investigators widely use in the healthcare industry, including for the submission of traditional Medicare FFS claims). While CMS has made the full transition to using MA diagnoses from encounter data when calculating MA risk scores, the CMS-HCC risk adjustment model calibrates for Part C on an FFS sample. CMS. "Report to Congress…." (2024), p. 64 (Accessed 03/06/2026)

[3] CMS. "Report to Congress: Risk Adjustment In Medicare Advantage. December 2024." (Accessed 03/06/2026)

[4] CMS uses the FFS United States Per Capita Cost (USPCC) to develop the benchmark for the annual rate book. (Accessed 03/06/2026)

[5] Medicare Payment Advisory Commission (MedPAC) provides a summary of the process. Reference "Chapter 4. Mandated report: Impact of changes in the 21st Century Cures Act to risk adjustment for Medicare Advantage enrollees." (Accessed 03/06/2026)

[6] Version 21 of the CMS–HCC model that CMS used before 2017 did not distinguish between beneficiaries with full versus partial dual Medicaid benefits. In 2017, CMS implemented a version of the CMS–HCC model (V22) to do this. It has six community segments: Non-dual aged, non-dual disabled, partial benefit dual aged, partial benefit dual disabled, full benefit dual aged, and full benefit dual disabled (MedPAC op cit, p.14)

models. In 2020, CMS implemented V24 model, with resolved risk scores reflecting a blend of the V22 and V24 models. CMS does not use all versions for payment. CMS only publishes and makes publicly available the models they use for payment.[7] For ESRD beneficiaries, CMS uses risk models that distinguish between beneficiaries on dialysis and beneficiaries that are post-transplant. These models have similar structures to the non-ESRD models and CMS calibrates to have an average risk score of 1.0 for these subpopulations. The CMS annual MA Rate Announcements document the applicable models for a payment year.[8]

The basic structure of each risk model relies on the same demographic factors and HCC structures, but the relative weights they assign to each factor differ. CMS assigns beneficiaries a "resolved" payment risk score based on the beneficiary's status in a payment month. In general, the status is a function of ESRD status, new enrollee status, and community or institutional status. CMS may also base risk score resolution on dual Medicare-Medicaid eligibility status. CMS adjusts model output risk scores for payment year specific FFS normalization and coding intensity adjustments. The FFS normalization adjustment accounts for changes in FFS population between the year the model is calibrated and the payment year. For Part C risk scores, the coding intensity adjustment accounts for changes in the diagnosis code submission patterns between FFS and MA. Both the normalization and the coding intensity factors can differ each year and researchers can find them in the HCC annual announcements on the CMS website.[9]

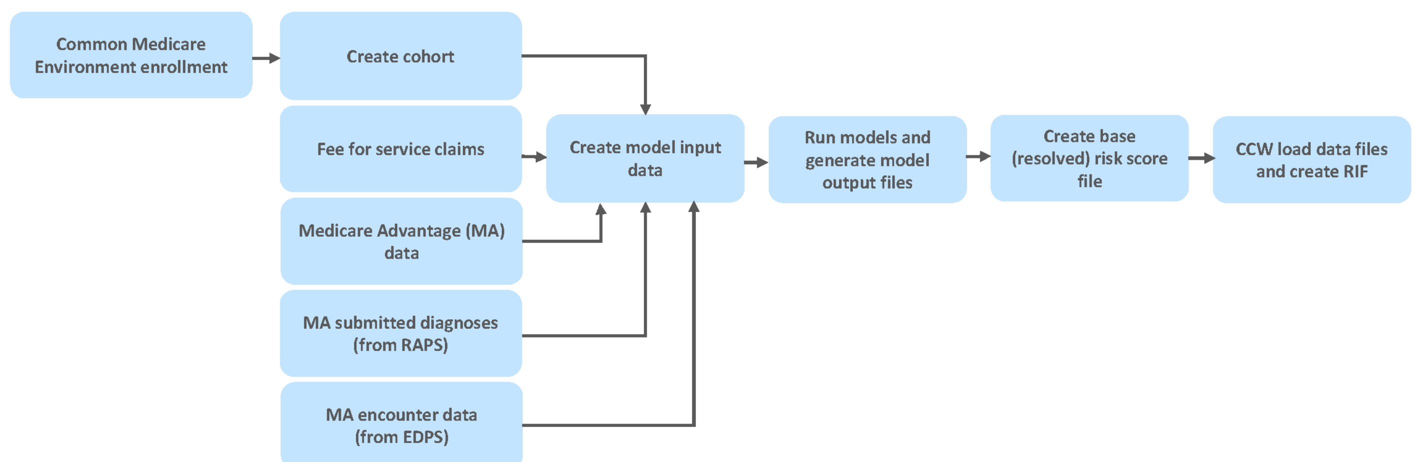
[7] CMS. "Report to Congress: Risk Adjustment In Medicare Advantage. December 2024." (Accessed 03/06/2026)
[8] CMS. Medicare Advantage Rates & Statistics. Announcements and Documents (updated each year) (Accessed 03/06/2026)
[9] CMS. Medicare Advantage Rates & Statistics. Announcements and Documents (updated each year) (Accessed 03/06/2026)

3.0 CCW Payment Risk Score Files

3.1 Source Data

CMS uses multiple sources to produce the Part C and Part D risk scores. These sources are the CMS Common Medicare Environment (CME) tables for the enrollment and demographic data, CMS FFS claims, MA plan-submitted diagnosis data (from the Risk Adjustment Processing System [RAPS]), and MA encounter data (from Encounter Data Processing System [EDPS]).[10] The final model runs that CMS produces for a given payment year use current beneficiary demographic information and diagnosis data for the data collection period submitted by the applicable submission deadline (i.e., January 2020 for payment year 2019, covering services provided in 2018). A visual depiction of the CMS data sources and processing to create the risk score files for researchers is in Figure 1.

Figure 1. Risk adjustment processing to create research risk score files



The risk adjustment models are prospective. The models use diagnosis codes from the year prior to the payment year to identify HCCs. For the 2019 CCW HCC data files, the payment/prediction year is 2019 and the model uses information from Medicare services provided in 2018.

CMS typically performs multiple operational model runs for each payment year, including an initial run, mid-year, and final. The CCW base (resolved) risk score files reflect the final run for each payment year. CMS runs the data for each Medicare beneficiary through each model and calculates risk scores for each model segment. Using these various MOFs, CMS creates a single record per beneficiary, with Part C and Part D risk scores that they normalize and adjust.

[10] CMS included encounter data in model runs starting in 2015 at various blend percentages (RAPS/FFS and EDPS/FFS inputs).

3.2 CMS HCC Data Files

CMS produces two main types of HCC files for each payment year, and makes the final files available to the CCW. Reference the CCW data dictionary for the record layouts.

- **The Risk Score Operational Payment base (resolved) file** — this is a beneficiary-level data file with variables containing monthly risk adjustment payment scores. CMS has normalized these monthly scores and applied a coding intensity factor where applicable. CMS has determined the appropriate Part C and Part D model segment risk score for each beneficiary/month and provides the corresponding risk adjustment payment score (i.e., a resolved risk score). The file also includes all possible Part C and Part D payment risk scores for reference purposes
- **MOFs** — Each MOF is a beneficiary-level file that contains the model-specific "raw" risk scores. MOFs provide beneficiary specific information including demographic characteristics and the set of diseases (HCCs) included in that particular model. Payment years may use a blend of model risk scores, and therefore uses different MOFs as inputs for CMS to calculate the resolved risk scores. Depending on the payment year, the detailed MOFs that are applicable may include (also reference Table 3):
    - ESRD V21 scores
    - V21 scores
    - V22 scores
    - V23 scores
    - V24 scores
    - RXHCC scores

The MOFs for an HCC model may vary by the array of variables used in the model and in some instances the relative coefficients assigned to each variable. Details about the applicable models and the calibration years are available in the CMS annual announcements.

4.0 CCW Payment Risk Score Files

The CCW team envisions that most investigators can conduct their research using only the base risk score file since all the applicable resolved beneficiary scores appear in this file. However, a brief description of the detailed MOFs follows. CMS and CCW do not provide technical assistance or user support to investigators who wish to use these files.

The CCW HCC files are linkable data files using the CCW BENE_ID.

4.1 CCW Risk Score Operational Payment Base File

The base (resolved) risk score file the CCW team provides to researchers uses a beneficiary's characteristics to identify the beneficiary's correct monthly payment score, i.e., the monthly "resolved" risk scores for each beneficiary.[11] There are four types of variables in the file: 1) overall monthly scores, 2) monthly model indicators, 3) long-term institutional monthly indicators, and 4) non-resolved payment scores. Reference Table 1 for a list of fields containing resolved scores. CMS calculates payment risk scores for all model segments for a beneficiary, regardless of the beneficiary's monthly characteristics.

In this document specific field names will be in all capitals.

Table 1. Base file — resolved payment risk scores

Variable name	Label
BENE_ID	CCW encrypted beneficiary ID
PTC_RISK_SCORE_01–12	Monthly Part C risk score 01 (January)–12 (December)
PTC_MODEL_SGMT_CD_01–12	Monthly Part C model segment code 01 (January)–12 (December)
PTD_RISK_SCORE_01–12	Monthly Part D risk score 01 (January)–12 (December)
PTD_MODEL_SGMT_CD_01–12	Monthly Part C model segment code 01 (January)–12 (December)

Investigators can use the monthly Part C and Part D model segment codes to identify the particular HCC model segment selected for that month. Table 2 lists the possible values for Part C model segment codes and Table 3 lists the possible values for the Part D model segment codes; reference the codebook on the CCW Data Dictionaries tab of the CCW website.

[11] It may not be possible to replicate the actual payment risk score for a plan using these files due to changes in the status of the beneficiary.

Table 2. Part C model segment code values, by payment year

PTC_MODEL_ SGMT_CD_MM	Description	2019	2020–2021
C	Community (Program of All-Inclusive Care for the Elderly [PACE] only 2017–2019)	X	—
C1	Community Post Graft 4–9 months (ESRD)	X	X
C2	Community Post Graft 10+ months (ESRD)	X	X
CF	Community Full Dual	X	X
CN	Community Non-Dual	X	X
CP	Community Partial Dual	X	X
D	Dialysis (ESRD)	X	X
E	New Enrollee	X	X
ED	New Enrollee Dialysis (ESRD)	X	X
E1	New Enrollee Post Graft 4–9 months (ESRD)	X	X
E2	New Enrollee Post Graft 10+ months (ESRD)	X	X
G1	Post Graft 1 month (ESRD)	X	X
G2	Post Graft 2–3 months (ESRD)	X	X
I	Institutional	X	X
I1	Institutional Post Graft 4–9 months (ESRD)	X	X
I2	Institutional Post Graft 10+ months (ESRD)	X	X
PA	PACE Dialysis	X	X
PB	PACE New Enrollee Dialysis	X	X
PC	PACE Community Post Graft 4–9 months (ESRD)	X	X
PD	PACE Institutional Post Graft 4–9 months (ESRD)	X	X
PE	PACE New Enrollee Post Graft 4–9 months (ESRD)	X	X
PF	PACE Community Post Graft 10+ months (ESRD)	X	X
PG	PACE Institutional Post Graft 10+ months (ESRD)	X	X
PH	PACE New enrollee Post Graft 10+ months (ESRD)	X	X
PI	PACE Community Partial Dual (starting 2020)	–	X
PJ	PACE Community Full Dual (starting 2020)	–	X
PK	PACE Community Non-Dual (starting 2020)	–	X
PL	PACE Post Graft 1 month (ESRD) (starting 2020)	–	X
PM	PACE Post Graft 2–3 months (ESRD) (starting 2020)	–	X
SE	New Enrollee Chronic Care Special Needs Plan (SNP)	X	X
Null/missing	Not applicable	X	X

Table 3. Part D model segment code values, by payment year

PTD_MODEL_ SGMT_CD_MM	Description	2019	2020–2021
D1	Community Non-Low Income Continuing Enrollee	X	X
D2	Community Low Income Continuing Enrollee	X	X
D3	Institutional Continuing Enrollee	X	X
D4	New Enrollee Community Non-Low Income Non-ESRD	X	X
D5	New Enrollee Community Non-Low Income ESRD	X	X
D6	New Enrollee Community Low Income Non-ESRD	X	X
D7	New Enrollee Community Low Income ESRD	X	X
D8	New Enrollee Institutional Non-ESRD	X	X
D9	New Enrollee Institutional ESRD	X	X
P1	PACE New Enrollee Community Low Income Non-ESRD	X	X
P2	PACE New Enrollee Community Non- Low Income Non-ESRD	X	X
P3	PACE New Enrollee Institutional Non-ESRD	X	X
P4	PACE New Enrollee Institutional ESRD	X	X
P5	PACE New Enrollee Community Low Income ESRD	X	X
P6	PACE New Enrollee Community Non- Low Income ESRD	X	X
P7	PACE Community Non- Low Income Continuing Enrollee	X	X
P8	PACE Community Low Income Continuing Enrollee	X	X
P9	PACE Institutional Continuing Enrollee	X	X
Null/missing	Not applicable	X	X

In addition to the Part C and Part D monthly model segment codes, the file also includes monthly long-term institutional (LTI) indicators (LTI_IND_01–12). The LTI indicators are consistent with the model segment codes when a beneficiary is a non-ESRD full risk beneficiary. The demographic variables that are in the base file are demographics that CMS used in the risk score development and within the risk score file; it is possible that some values may not exactly match values in the CCW Medicare Beneficiary Summary File (MBSF) due to timing differences. Additional details regarding the MBSF appear in section 6.0 Medicare Enrollment Data.

The largest proportion of individuals live in the community setting. Individuals in long-term care facilities have different medical needs and different costs. PACE organizations are a distinct set of contracts that target specific Medicare populations and are a relatively small proportion of the MA enrollment.

As with the Medicare Part C HCCs, CMS has developed several Part D model segments (e.g., Rx model segments). These include: 1) aged, 2) disabled, 3) institutional, and 4) new enrollee. CMS further divides some of these segments into low income and not low income. As with Part C risk scores, CMS calculates Rx risk scores using FFS claims and MA reported diagnosis information to predict expenditures associated with prescription drugs. Additional details are available on the CMS website.[12,13]

A listing of all variables included in the data file appears in the CCW Data Dictionaries tab of the CCW website.

It is possible that some investigators may wish to calculate a single annual payment risk score. If this is desirable, then CCW recommends averaging the monthly scores to calculate the annual score.

[12] CMS. Annual "Rate Announcement and Call Letters" are available for Prescription Drug Coverage Contractors. For example: (Accessed 03/06/2026)
[13] CMS. Annual "Rate Announcement and Call Letters" are available for Prescription Drug Coverage Contractors. For example: (Accessed 03/06/2026)

4.2 CCW Detailed MOFs

The MOFs that CMS uses as inputs for the resolved risk scores vary by year, in accordance with CMS payment model requirements. The CMS annual Rate Announcements document these requirements.[14]

The MOFs for each of the prospective risk adjustment models include only the HCCs that best predict future (next year's, i.e., the payment year [e.g. the 2019 CCW files]) Medicare Part A and Part B expenditures. CMS updates model versions for a variety of reasons, including changes in valid diagnoses mappings, updates to the years used to calibrate the model, and changes in clinical and policy objectives.[15] The annual advance notice includes proposed models that CMS may implement for a given payment year. The annual announcement specifies which specific models CMS is using for the payment year.[16]

Table 4. HCC model segments by payment year

Files	2019	2020	2021
Base	X	X	X
V21	X	—	—
V22	X	X	X
V23	X	—	—
V24	—	X	X
ESRD V21	X	X	X
RX	X	X	X

For 2019 payment year, there are five MOFs (Table 4). The MOFs are model specific and contain the risk scores that CMS assigns to a beneficiary for payment. For example, the ESRD model contains the various segments that CMS uses to calculate the risk score for the set of ESRD statuses assigned to a beneficiary: dialysis, transplant, and post-graft. CMS uses the V21 model for PACE enrollees; and the V22/V23 and V22/V24 models to create a blended risk score for MA enrollees.

Starting with the V23 Risk Score Model, CMS began implementing changes to address the legal requirements of the 21st Century Cures Act.[17] CMS designed the updated Risk Score Models to account for the presence of mental health and substance use disorders, and severity of chronic kidney disease, and starting with the V24 Risk Score Model, the total number of diseases or conditions of an individual.

CCW provides unique data file variable names for each variable in the MOFs, so it is easy to trace which version of the HCCs is in the data. If a beneficiary has more than one condition within a particular HCC hierarchy, CMS uses only the most severe condition when determining his or her risk score. Each condition variable has a value of 1 if the beneficiary had the condition in the prior year; otherwise, the value is 0.

4.2.1 ESRD V21 Scores

Individuals who experience ESRD have high medical costs and more extensive medical needs than the general Medicare population; therefore, this group of individuals has their own HCC model. The ESRD HCC model divides the population into four segments: 1) those undergoing dialysis, 2) those who have received kidney transplants (from the

[14] CMS. Medicare Advantage Rates & Statistics. Announcements and Documents (updated each year) (Accessed 03/06/2026)
[15] Pope GC, Kautter J, Ingber MJ, et al. "Evaluation of the CMS-HCC Risk Adjustment Model." Final Report to CMS. March 2011. (Accessed 03/06/2026)
[16] The Annual Medicare Advantage and Part D Advance Notice can be found here: and select the year of "Advance Notice" of interest. (Accessed 03/06/2026)
[17] CMS. "CMS Risk Adjustment Model Research Findings." October 2018. (Accessed 03/06/2026)

time of transplant to three months post graft), 3) those four to nine months after a transplant (a functioning graft), and 4) those who are 10 months or more after a transplant. This file contains the same HCC conditions as the V21 file

CMS produces three versions of the ESRD file each year using various data inputs for the model; therefore for each beneficiary there are three records in the annual ESRD file. CCW derived a field called the MOF diagnosis input file source code (MOF_DGNS_INPUT_CD) in the MOFs to identify the three versions. Table 5 describes the field values.

Table 5. Identified versions of the MOF diagnosis input file source code

Valid values	Value descriptions
PPME	Encounter, FFS, RAPS for IP (varies by year)
PPMER	Encounter, FFS, RAPS (this is the PACE path)
PPMR	RAPS and FFS (consistent longitudinally)

The first three letters of the values (the “PPM”) mean “post-process model.” CMS runs the HCC files through this final processing cycle before CCW loads the data; therefore, this portion of the value does not vary by MOF or over time.

4.2.2 V21 Scores

CMS developed the HCC model called version 21 (V21) in 2012. CMS includes the V21 model scores for the 2019 payment year only (no other year of HCC data that CCW delivers). In the 2019 risk score data files, CMS used this model for PACE enrollees. Compared to earlier HCC model versions, this model incorporated additional HCC conditions, different disease interactions, and different disabled HCC interactions terms to calculate an HCC score.[18] Nearly one-third of the variables in the file are variables for HCC conditions. All the HCCs within the V21 file have a variable name with “V21_” prefix (Table 6 is an example, HCC1 (HIV/AIDS) has the name V21_HCC1 in this data file (whereas the same HCC in the V22 file is called V22_HCC1). The HCC numbering may apply to different condition categories in different models. For example, in V21 and V22 HCC54 was Drug/Alcohol Psychosis; for V23 and V24 HCC54 is Substance Use with Psychotic Complications.

Table 6. Naming conventions for HCCs — V21 example

Variable name	Condition
V21_HCC1	HIV/AIDS (V21 HCC1)
V21_HCC2	Septicemia/shock (V21 HCC2)
V21_HCC6	Opportunistic infections (V21 HCC6)
V21_HCC8	Metastatic cancer/acute leukemia (V21 HCC8)

Each condition variable has a value of 1 if the beneficiary had the condition in the prior year; otherwise, the value is 0.

There are also variables to indicate some disease/disease combinations, and some disability/disease combinations. Demographic variables that are related to the segment scores (e.g., sex, age groupings, and original reason for enrollment in Medicare) are also available.

4.2.3 V22 Scores

[18] Pope GC, Kautter J, Ingber MJ, et al. “Evaluation of the CMS-HCC Risk Adjustment Model.” Final Report to CMS. March 2011. (Accessed 03/06/2026)

CMS developed the V22 segmented model in 2017. Approximately one-third of the variables in the file represent distinct condition categories in V22 of the HCC risk-adjustment model,[19] CMS then groups some into disease hierarchies. The variable names for each of these HCCs begins with "V22_".

This file is available starting with 2019. Starting with 2020, CMS produces two versions of the V22 file per year using various data inputs for the model; therefore for each beneficiary there are two records in the annual V22 file. CCW derived a field called the MOF diagnosis input file source code (MOF_DGNS_INPUT_CD) in the MOFs to identify the two versions. Table 5 describes the field values. The two values for V22 are MOF_DGNS_INPUT_CD = PPMER and PPMR.

4.2.4 V23 Scores

CMS includes the V23 model scores for the 2019 payment year only (no other year of HCC data that CCW delivers). They include distinct HCC condition categories, as well as disease hierarchies. This model is largely V22, but CMS modified it to add new HCCs for substance abuse, mental health, and stage three chronic kidney disease (CKD).[20]

4.2.5 V24 Scores

CMS began using the V24 model in 2020. It included additional categories for the number of conditions for each beneficiary.[21]

4.2.6 RXHCC Scores

CMS uses the RxHCC model to estimate Medicare Part D prescription drug plan liability costs. As a result, there are similar characteristics as the other HCC models. For each beneficiary, CMS uses diagnosis codes to calculate RxHCC categories; however, they use a different HCCs to calculate a risk score than they use for the other Part C HCCs.

In the RxHCC model, nearly one-half of the variables are the HCC condition categories included in the calculation of a risk score. There are also variables for select disease and disabled interactions, as well as post graft add-on factors applicable to beneficiaries with ESRD who have received a transplant and are in post graft status.[22]

The conditions included in the model are clinically meaningful regarding expected prescription drug expenditures. For example, CMS includes hypertension in the RxHCC score calculation but not in the Part C HCC models. The exact number of variables in the RxHCC file depends on the payment year.

CMS produces three versions of the RXHCC file each year using various data inputs for the model; therefore for each beneficiary there are three records in the annual RXHCC file. CCW derived a field called the MOF diagnosis input file source code (MOF_DGNS_INPUT_CD) in the MOFs to identify the three versions. Table 5 describes the field values.

The CCW base risk score file contains the resolved Part D payment risk scores.

---

[19] CMS. "HHS-Operated Risk Adjustment Methodology Meeting. Discussion Paper." March 24, 2016. (Accessed 03/06/2026)
[20] MedPAC, "Chapter 4. Mandated report: Impact of changes in the 21st Century Cures Act to risk adjustment for Medicare Advantage enrollees" (p. 104) (Accessed 03/06/2026)
[21] MedPAC (op cit)
[22] CMS Risk Adjustment Methodology, December 2021. (p. 2–14) (Accessed 03/06/2026)

5.0 Medicare Risk Scores

5.1 Calculation of Risk Scores

The purpose of the Medicare risk scores is to estimate a relative cost factor. CMS calculates individual beneficiary-level risk scores by adding the relative factors associated with each beneficiary's demographic and disease factors. CMS sums the weight/coefficient for each applicable factor for the beneficiary for the applicable model segment.

The risk adjustment model is based on a clinical classification system that maps the International Classification of Diseases (ICD) diagnosis codes into a system of diagnostic categories (the HCCs).[23] The conditions are hierarchical meaning that CMS selects the most severe condition within a category and accumulates scores from all unrelated conditions categories. Some combinations of clinical conditions are important and the effect of the combination is not just additive (i.e., HCC interaction terms).

Table 7 and Table 8 show the categories of data elements that CMS uses in determining both the raw and the resolved risk scores.

Table 7. Variables added together to create the raw risk scores within the risk adjustment model

Input variables	Details
Demographic variables: • Age/sex • Originally entitled to Medicare due to disability • Medicaid	There are relative factors associated with each demographic variable.
Other payment information:[24] • ESRD status • LTI status • Dual enrollment status • PACE frailty score*	CMS uses separate risk adjustment models for ESRD beneficiaries, PACE beneficiaries, and non-ESRD and non-PACE enrolled beneficiaries.
Disease variables: • Disease HCCs • Disease/Disabled interactions	CMS uses diagnoses submitted by plans and FFS providers to assign HCCs and interactions. There are relative factors associated with each HCC and interaction.
Sum of factors Demographic + Disease = Raw risk scores	CMS adds the relative factors for all the demographic variables, HCCs, and interactions together. The result is the raw risk score.

* CMS only adds frailty adjustments to the risk scores for PACE participants and community-based beneficiaries enrolled in qualifying Fully Integrated Dual Eligible Special Needs Plans (FIDE SNPs). For FIDE SNPs to qualify for a frailty adjustment, they must have a similar level of frailty as PACE organizations. There is no frailty adjustment applied to the risk scores of long-term institutionalized beneficiaries or ESRD beneficiaries.[25]

[23] CMS. "HHS-Operated Risk Adjustment Methodology Meeting. Discussion Paper." March 24, 2016. (Accessed 03/06/2026)
[24] CMS Risk Adjustment Methodology, December 2021. (Accessed 03/06/2026)
[25]CMS Risk Adjustment Methodology, December 202. (Accessed 03/06/2026)

Table 8. Adjustments within the risk adjustment model

Adjustment	Details
Payment adjustments	CMS applies a *normalization* factor to keep the average FFS risk score at 1.0.
Payment adjustments **NOTE**: CMS applies this coding intensity adjustment to all scores (not limited to scores for MA enrollees).	CMS applies a *coding pattern adjustment* to account for differential coding patterns between MA plans and FFS.
Final step	Resolve the risk score for each beneficiary.

CMS applies adjustments to the raw risk scores to convert them to payment scores, reference Table 8.

- **Normalization** — CMS uses the normalization factor to keep the FFS average risk score to be equal to 1.0 for the payment year. Due to the lag between the time CMS estimated the model data and converted it to a relative risk score, and the payment year, changes in the demographic composition and health conditions of the FFS cohort may result in an average risk score not equal to 1.0. CMS designed the normalization factor to correct for this situation

- **Coding intensity** — CMS designed the adjustment for MA coding pattern differences (i.e., the MA coding intensity adjustment) to account for the relative difference in reporting of diagnosis codes between the FFS population used to estimate the models and the MA cohort that CMS is reimbursing based on the model risk score

  NOTE: CMS applies the coding intensity adjustment to the resolved risk scores for all beneficiaries regardless of payment year enrollment status. CMS uses the payment risk scores to determine MA reimbursement. As a result, the resolved risk score for a beneficiary that was FFS in the data collection period may be artificially low as the coding intensity adjustment may not be applicable. Investigators may "back-out" this coding intensity adjustment by identifying FFS beneficiaries (i.e., using the CCW MBSF) and dividing the resolved risk scores by 0.9410. The actual adjustment factor may vary by year and CMS publishes it in the CMS Annual Rate Announcement and Call Letter for MA plans.

- **Resolved scores** — the last step in the risk adjustment process is to determine which of the beneficiary's converted payment scores applies to the beneficiary for a given month. The resolved risk score multiplied by the applicable risk rate is what determines the payments made to the MA plan

The base (resolved) risk score file has each beneficiary's monthly Part C and Part D "resolved" risk scores (i.e., the PTC_RISK_SCORE_MM variable and PTD_RISK_SCORE_MM, where MM is the month 01 through 12), and the associated model segment code indicator (the risk score factor code, or PTC_MODEL_SGMT_CD_MM and PTD_MODEL_SGMT_CD_MM) that indicates which segment is appliable for that month. The base file also contains the additional relative risk scores associated with the beneficiary for both Part C and Part D.

5.2 Hierarchical Condition Categories

The CMS-HCC diagnostic classification system begins by classifying all ICD-10-CM diagnosis codes into diagnostic groups. A diagnostic group represents a well-specified medical condition or set of conditions. CMS further aggregated the diagnostic groupings into Condition Categories (CCs). CCs describe a broader set of similar diseases. Although they are not as homogeneous as the diagnostic groupings, diseases within a CC are related clinically and with respect to cost. An example is the CCs for diabetes, where both the diabetes with acute complications and diabetes with chronic complications include diagnosis codes for ketoacidosis or coma. Note, a small number of diagnosis codes mapped into more than one CC.

CMS imposes hierarchies among related CCs, so that the model codes a person for only the most severe manifestation among related diseases. After imposing hierarchies, CCs become HCCs. For example, the model organizes the diabetes diagnosis codes in the diabetes hierarchy, consisting of three CCs arranged in descending order of clinical severity and

cost, from 1) diabetes with acute complications to 2) diabetes with chronic complications to 3) diabetes without complication. Thus, a person with a diagnosis code of diabetes with acute complications, diabetes with chronic complications, and diabetes without complication, with diagnosis code of diabetes with acute complications. Although HCCs reflect hierarchies among related disease categories, for unrelated diseases, the model accumulates the relative coefficient for each HCC, i.e., the model is "additive." For example, a female with both rheumatoid arthritis and breast cancer has (at least) two separate HCCs coded, and her predicted cost reflects increments for both conditions.

Because the model codes a single individual for none, one, or more than one HCC, the CMS-HCC model can individually price tens of thousands of distinct clinical profiles. The model's structure thus provides, and predicts from, a detailed comprehensive clinical profile for each individual. An evaluation of the CMS risk models is publicly available on the CMS website.[26]

[26] Pope GC, Kautter J, Ingber MJ, et al. "Evaluation of the CMS-HCC Risk Adjustment Model." Final Report to CMS. March 2011. (Accessed 03/06/2026)

6.0 Medicare Enrollment Data

Investigators using the CCW risk score files may want to use information from Medicare enrollment files to identify and describe the study population. Data for 100% of Medicare-enrolled beneficiaries is available from the CCW. The CCW updates the annual enrollment data files known as the MBSF using downloads from the CMS CME database. Each annual file includes all beneficiaries documented as being alive for some part of the calendar year and enrolled in the Medicare program (Part A or Part B) for at least one month of the year.

CCW removes the Medicare beneficiary health insurance claim (HIC)[27] numbers from all data files delivered to researchers. CCW adds a unique CCW beneficiary identifier (variable called the BENE_ID) in each data file delivered as part of the output package. The unique CCW beneficiary identifier provides a common link across all applicable types of data available, thus allowing data users to link the CCW Payment Risk Score data to beneficiary and claims data in the CCW.

The unique CCW beneficiary identifier field is specific to the CCW and is not applicable to any other identification system or data source. CCW encrypts this identifier prior to delivering the data files to researchers. In addition, CCW encrypts all data files delivered to researchers (details in 8.0 Further Assistance with CCW Data). Each research request employs a different encryption key for the beneficiary identifier field and the data files.

The record layouts for all the CCW data files are under the Data Dictionaries tab on the ccwdata.org website.

The MBSF contains information on beneficiaries' demographic characteristics and details of their enrollment in Medicare.

Examples of the types of information in the MBSF include:

- Demographic — age, race, sex, date of death
- Geographic — state, county, zip code
- Enrollment — the start date for Medicare coverage; the beneficiary Medicare entitlement reason(both the original reason and the current reason, which can differ); and monthly information on eligibility (Part A, Part B, or both), enrollment in MA (aka Medicare Part C), and enrollment in Part D

The base beneficiary summary file is also known as the MBSF Part A, B, C, and D segment.[28]

Although there is a small amount of beneficiary demographic information on the detailed MOFs, in general the CCW team recommends that researchers use the MBSF to obtain these data. CMS creates the MOFs at a different point in time, during which actual values (e.g., age or disability status) may have changed, and in rare instances, the model uses different decision rules for populating fields than the MBSF. Investigators should consider the MBSF the definitive source of Medicare beneficiary enrollment and demographic information.

[27] CMS began using a new Medicare beneficiary identifier (MBI) in place of the HIC, starting in 2018.
[28] Beginning with the 2023 annual data file, there are additional fields and adjustments in algorithms for the MBSF; therefore, the CCW team refers to the updated data file as "version 2" of the MBSF ABCD.

7.0 Format, Content, and Encryption of CCW Output Files

This section describes the content and format of the CCW payment risk score output package. CCW delivers this data physically to researchers or they may use it within the CCW Virtual Research Data Center (VRDC).

The CCW team packages files delivered physically in encrypted self-decrypting archives (SDAs). The encryption technique for files extracted from the CCW uses Pretty Good Privacy Command Line 9.0 with the SDA method. This method builds a compressed, encrypted, password protected file using a FIPS 140-1/140-2 approved AES256 cipher algorithm. The CCW team builds the SDA on the CCW production server, downloads it to a desktop PC, and burns it to a CD, DVD, or USB hard drive depending on the size of the files.

Table 9 lists the format and descriptions of each of these SDA contents. Table 10 lists the file names of the data available within the CCW VRDC.

Table 9. Format and naming convention for the CCW physically delivered files

File	File description
README_FIRST_REQ232_2019.txt	This is a text file that describes the files contained in the output package. Filename example: readme_first_req232_2019.txt
res000011111req000232_2019_RSBASE.exe	This is the executable that researchers must run to decrypt and uncompress the data file. In this example, 000011111 is the DUA number, 232 is the request number, and 2019 is the year of the data. This executable includes v6 and v8 SAS read-in programs, the .dat file, and .fts file.
risk_score_base_res000011111_req000232_2019.dat	This is the data file contained within the executable file.
risk_score_base_res000011111_req000232_2019.fts	This is a file transfer summary document, found within the SDA that shows the file layout and record counts.
risk_score_v21_read_v6.sas risk_score_v21_read_v8.sas	This is the v6 and v8 SAS read-in programs found within the SDA.

Detailed MOFs

File	File description
res000011111req000232_2019_RSV21.exe risk_score_v21_res000011111_req000232_2019.dat risk_score_v21_res000011111_req000232_2019.fts risk_score_v21_read_v6.sas risk_score_v21_read_v8.sas	This set of files includes the V21 MOF executable file, and once decrypted and uncompressed it includes the .dat (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
res000011111req000232_2019_RSV22.exe risk_score_v22_res000011111_req000232_2019.dat risk_score_v22_res000011111_req000232_2019.fts risk_score_v22_read_v6.sas risk_score_v22_read_v8.sas	This set of files includes the V22 MOF executable file, and once decrypted and uncompressed it includes the .dat (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
res000011111req000232_2019_RSV23.exe risk_score_v23_res000011111_req000232_2019.dat risk_score_v23_res000011111_req000232_2019.fts risk_score_v23_read_v6.sas risk_score_v23_read_v8.sas	This set of files includes the V23 MOF executable file, and once decrypted and uncompressed it includes the .dat (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.

File	File description
res000011111req000232_2019_RSV24.exe* risk_score_v24_res000011111_req000232_2019.dat risk_score_v24_res000011111_req000232_2019.fts risk_score_v24_read_v6.sas risk_score_v24_read_v8.sas	This set of files includes the V24 MOF executable file, and once decrypted and uncompressed it includes the .dat (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
res000011111req000232_2019_RSESRD.exe risk_score_esrd_res000011111_req000232_2019.dat risk_score_esrd_res000011111_req000232_2019.fts risk_score_esrd_read_v6.sas risk_score_esrd_read_v8.sas	This set of files includes the ESRD MOF executable file, and once decrypted and uncompressed it includes the .dat (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
res000011111req000232_2019_RSRX.exe risk_score_rx_res000011111_req000232_2019.dat risk_score_rx_res000011111_req000232_2019.fts risk_score_rx_read_v6.sas risk_score_rx_read_v8.sas	This set of files includes the RX MOF executable file, and once decrypted and uncompressed it includes the .dat (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.

*V24 is not applicable for 2019; the CCW team includes this only as an example of the naming convention.

Different MOFs apply for each payment year; they include only MOFs applicable to the years in the data request. The files within the CCW VRDC have a different naming convention, as depicted in Table 10.

Table 10. Payment risk score files in the CCW VRDC

RISK_SCORE_BASE_YYYY	This is name of the base risk score file in the CCW VRDC.
RISK_SCORE_V21_YYYY	This is name of the V21 MOF risk score file in the CCW VRDC.
RISK_SCORE_V22_YYYY	This is name of the V22 MOF risk score file in the CCW VRDC.
RISK_SCORE_V23_YYYY	This is name of the V23 MOF risk score file in the CCW VRDC.
RISK_SCORE_V24_YYYY	This is name of the V24 MOF risk score file in the CCW VRDC.
RISK_SCORE_ESRD_YYYY	This is name of the ESRD MOF risk score file in the CCW VRDC.
RISK_SCORE_RX_YYYY	This is name of the RX MOF risk score file in the CCW VRDC.

8.0 Further Assistance with CCW Data

Researchers interested in working with CCW data should contact ResDAC. ResDAC is a CMS contractor and researchers should first submit requests to ResDAC for assistance in the application, obtaining, or using the CCW data. They offer free assistance to researchers using Medicare data for research. The ResDAC website 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 Medicare data, and other helpful resources.

If a ResDAC technical advisor is not able to answer researchers' questions, the technical advisor directs the researcher to the appropriate person. If researchers require additional CMS data (data not available from the CCW) to meet research objectives, or the researcher has any questions about other data sources, the researcher should first visit the ResDAC website.

The CCW Help Desk provides assistance between 8:00 am to 5:00 pm ET, Monday through Friday (excluding most federal holidays). Contact the CCW Help Desk at ccwhelp@ccwdata.org or 1-866-766-1915.

Appendix A — Acronyms

Acronym	Definition
CC	Condition category
CCW	Chronic Conditions Warehouse
CME	Common Medicare Environment
CMS	Centers for Medicare & Medicaid Services
DUA	Data Use Agreement
EDPS	Encounter Data Processing System
ESRD	End-stage renal disease
FFS	Fee-for-service
HCC	Hierarchical Condition Category
FIDE SNP	Fully Integrated Dual Eligible Special Needs Plan
HIC	Health Insurance Claim number
ICD	International Classification of Diseases
LTI	Long-term institutional
MA	Medicare Advantage
MA-PD	Medicare Advantage Prescription Drug
MBSF	Master Beneficiary Summary File
MedPAC	Medicare Payment Advisory Commission
MOF	Model output file
PACE	Program of All-Inclusive Care for the Elderly
PDE	Prescription drug event
PPM	Post-process model
RAF	Risk adjustment factor
RAPS	Risk Adjustment Processing System
ResDAC	Research Data Assistance Center
RIF	Research Identifiable File
SDA	Self-decrypting archive
SNP	Special needs plan
USPCC	United States Per Capita Cost
VRDC	Virtual Research Data Center