Chronic Conditions Warehouse

Your source for national CMS Medicare and Medicaid research data

Chronic Conditions Warehouse Virtual Research Data Center

Medicare Oncology Care Model (OCM) Data Files User Guide

MAY 2024 | VERSION 1.0

Revision Log

Date	Changed by	Revisions	Version
May 2024	K. Schneider	Created initial document	1.0

Table of Contents

List of Tables

1.0 Overview

The Center for Medicare & Medicaid Innovation (CMS Innovation Center) develops new payment and delivery models designed to improve the effectiveness and efficiency of specialty care. One of these specialty models was the Oncology Care Model (OCM) that operated from July 2016 through June 2022, and aimed to provide higher quality, coordinated oncology care at the same or lower cost to Medicare. Under OCM, physician practices entered into payment arrangements with CMS that included financial and performance accountability for episodes of care surrounding chemotherapy administration to cancer patients. The Centers for Medicare & Medicaid Services (CMS) also partnered with commercial payers in the model. The participants in OCM committed to providing enhanced services to Medicare beneficiaries such as care coordination, navigation, and adherence to national treatment guidelines for care.

CMS uses the Chronic Conditions Warehouse (CCW) to develop and manage CMS research data resources. The CCW has complete (100%) Medicare enrollment and fee-for-service (FFS) claims data, obtained directly from CMS. CCW acquired OCM data files from CMS. From this source data, the CCW team has prepared data files to disseminate to researchers and certain government agencies that CMS has approved under a Data Use Agreement (DUA) to obtain OCM data for research purposes. The CCW OCM clinical and staging Medicare data files contain identifiable information. They are subject to the Privacy Act and other federal government rules and regulations (reference the Research Data Assistance Center [ResDAC] website for details on requesting Medicare data http://www.resdac.org/).

This guide provides researchers with information to clarify their work with OCM data files. Appendix A lists abbreviations used in this document.

2.0 Background

The CMS Innovation Center conducted OCM model from July 2016 through June 2022. OCM's original end date was June 2021; however, CMS extended the model for one year due to the COVID-19 public health emergency (PHE). The goal of OCM was to utilize appropriately aligned financial incentives to improve care coordination, appropriateness of care, and access to care for beneficiaries undergoing chemotherapy. OCM encouraged participating practices to improve care and lower costs through an episode-based payment model that financially incentivized high-quality, coordinated care. The CMS Innovation Center expected that these improvements would result in better care, smarter spending, and healthier people. CMS expected practitioners in OCM to rely on the most current medical evidence and shared decision-making with beneficiaries to inform their recommendation about whether a beneficiary should receive chemotherapy treatment. OCM provided an incentive to participating physician's practices to comprehensively and appropriately address the complex care needs of the beneficiary population receiving chemotherapy treatment, and heighten the focus on furnishing services that specifically improve the patient's experience or health outcomes.

OCM incorporated a two-part payment system for participating practices, creating incentives to improve the quality of care and furnish enhanced services for beneficiaries who undergo chemotherapy treatment for a cancer diagnosis. The two forms of payment included a per-beneficiary-per-month monthly enhanced oncology services (MEOS) payment and the potential for a performance-based payment (PBP) for six-month episodes of chemotherapy care. The $160 MEOS payment assisted participating practices in effectively managing and coordinating care for oncology patients during episodes of care, while the potential for PBP incentivized practices to lower the total cost of care and improve care for beneficiaries during treatment episodes. The PBP incorporated either one-sided or two-sided risk. Participants electing two-sided risk were eligible to earn higher PBPs, but were also at-risk of owing recoupments. For episodes initiated July 2019 through December 2021 (corresponding with the seventh through eleventh performance periods), CMS gave OCM participants the option to "opt out" of reconciliation (neither earn a PBP, nor owe a recoupment) as part of a series of flexibilities offered to support participants during the COVID-19 PHE.

OCM launched with 176 participants across the country. By July 2019, there were 157 participants, and by January 2020 there were 104, almost all of whom remained in the model until its end in June 2022. The decrease in participation from July 2019 to January 2020 reflects a model requirement that participants not earning a PBP through the fourth performance period either had to elect two-sided risk or exit the model.

CMS documented additional details related to the model design, as well as annual evaluation reports pertaining to the model, on the CMS Innovation Center OCM website.

2.1 OCM Population

OCM included episodes initiated on the date of service for an initial Part B chemotherapy drug claim with a corresponding cancer diagnosis on the claim, or on the fill date for an initial Part D chemotherapy drug claim with a corresponding Part B claim for cancer on the date of, or in the 59 days preceding, the drug claim. Episodes continued for six months after the triggering chemotherapy event, as long as the beneficiary met eligibility requirements for all six months. Beneficiaries who continued to receive chemotherapy after completing the six-month episode initiated a new episode. CMS organized the episodes by performance periods. Performance periods are the six-month periods of time during which a cohort of episodes terminated and was reconciled together.

CMS assigned episodes to one of the cancer types included in the model based on the plurality of diagnoses on the evaluation and management (E&M) visits for cancer that occurred during the six months of the episode. This list of diagnosis codes is available on the OCM website here:
https://www.cms.gov/priorities/innovation/media/document/ocm-cancer-type-mapping-codes-pp1and2-xls

CMS attributed each six-month episode included in the OCM datasets to the OCM practice associated with the most qualifying E&M visits for cancer care during the six-month episode. The Healthcare Common Procedure Coding System (HCPCS), code ranges 99201–99205 and 99211–99215, defined E&M visits that qualified for inclusion.

2.2 OCM Data Collection/Registry

OCM participants submitted staging and clinical data to the OCM Data Registry (OCMR), a web-based data submission and collection tool, for episodes the model attributed to their practice for each performance period, or each six-month cohort of episodes. Participants provided beneficiary data by manual input or by uploading the OCMR staging abstraction tool or the OCMR staging upload template populated with staging, clinical, and demographic data.

Every six months, CMS copied and exported the beneficiary registry data to the CCW environment for the OCM implementation contractor's use. This allowed the implementation contractor to perform required analysis on the exported OCMR data.

CMS required OCM practices to report clinical and staging data associated with each beneficiary's assigned cancer. Some of these data elements were used to risk-adjust the financial targets that determined performance in the model. In the first five performance periods, if a practice did not report to the OCMR, CMS did not consider it eligible for a performance-based payment. For the sixth through eleventh performance periods, CMS permitted practices to opt out of reporting due to external burdens associated with the COVID-19 PHE, and were still eligible to receive performance-based payments.

2.3 Project Data Files

The data from OCM is a set of two linkable data files, from three different sources: 1) information the participants submitted into the OCMR, 2) information CMS obtained from the CCW Master Beneficiary Summary File (MBSF), and 3) information CMS derived. The CCW OCM codebook identifies each variable in the data sets as originating from one of these three sources. The data includes information about the episodes, the attributed practice ID, episode beginning and ending dates, episode cancer types, and the clinical and staging characteristics that CMS required practices to report. Though the files contain dates of birth and death, as well as each beneficiary's sex, they do not contain names, beneficiary street addresses, or contact information. The files include data reported for the third through eleventh performance periods, or, all episodes initiated between July 2, 2017, and December 31, 2021. CMS has not shared the first two performance periods of data because CMS encountered obstacles that made the data reported questionable in accuracy and reliability.

The first file is the OCM clinical and staging data file. This file contains information about each episode initiated between July 2, 2017, and December 31, 2021 (third through eleventh performance periods) and the information practices reported to the OCMR. There is one row per episode, meaning that unique beneficiaries can appear multiple times in the file depending on the number of six-month episodes initiated.

The second file is the OCM current clinical status file. This file is linkable to the clinical and staging data file, and includes information on the reported current clinical status (CCS). CMS required that practices report CCS at least once for each episode, as well as whenever there was a change in the CCS. This means that many beneficiaries have multiple CCS records, either per episode or across multiple episodes. The OCM team associates a current clinical status date (CRNT_CLNCL_STUS_DT) with each CCS record.

3.0 CCW OCM Data Files

CMS uses the CCW to develop and manage CMS research data resources. The CCW team obtains the OCM data files from CMS and disseminates them. The CCW BENE_ID allows for linkage to other CCW data products (e.g., Medicare enrollment and claims). Throughout this section, the CCW team writes SAS variable names in all capital letters.

3.1 Clinical and Staging Data

The OCMR includes all episodes eligible for OCM reconciliation in the third through eleventh performance periods in the OCM clinical and staging data. CMS required practices to report on only 75% of those episodes to be eligible for a PBP (unless they opted out of reporting in performance periods six through eleven). Therefore, not all episodes in the file will have OCMR information populated (Table 1). There are blank values in the file variables sourced from the OCMR when a practice did not report the episode. The following information is available in the clinical and staging data:

- Beneficiary level — unique identification number, date of birth, sex, date of death
- Episode level — episode beginning date, episode ending date, attributed practice ID, cancer type (version that OCM has reconciled), performance period, indication of whether the practice reported data for the episode
- For the episodes where practices reported data, the clinical and staging data includes reported ICD-10 diagnosis, initial diagnosis date, cancer type (reported by practice), AJCC edition, primary tumor-nodal disease-metastasis (TNM) values, estrogen receptor, progesterone receptor, HER2 amplification, and histology. Not all clinical and staging data elements apply to all cancer types. The codebook provides additional details.

Table 1. OCM clinical and staging file variables

Types of data	Long SAS name	Label
Beneficiary level	BENE_ID	Encrypted CCW beneficiary ID
	DOB	Beneficiary date of birth
	SEX	Beneficiary sex
	DOD	Beneficiary date of death
Episode level	EPSD_BGN_DT	Episode beginning date
	EPSD_END_DT	Episode ending date
	OCM_PRCTC_ID	OCM-assigned practice ID
	CNCR_TYPE_ASGNED	OCM-assigned cancer type
	PRFMNC_PRD	OCM six-month episode performance period
	RPTD_TO_RGSTRY	Participant reported episode to OCMR
Subset of episodes (clinical and staging data)*	ICD10_CNCR_DGNS_CD_RPTD	ICD-10 cancer diagnosis code reported by participant
	CNCR_TYPE_RPTD	Reported cancer type
	INITL_DGNS_DT	Initial diagnosis date for episode
	AJCC_EDITION	AJCC cancer staging edition
	PRMRY_TUMOR	Primary tumor staging
	NODAL_DISEASE	Nodal disease staging
	METASTASIS	Metastasis staging
	ESTROGEN_RECEPTOR	Breast cancer estrogen receptor
	PROGESTERONE_RECEPTOR	Breast cancer progesterone receptor
	HER2_AMPLIFICATION	HER2 amplification
	HISTOLOGY	Histology

* Some staging variables apply only to a subset of cancer types. The codebook has additional details.

3.2 Current Clinical Status (CCS) Data

The CCS data file contains records only for reported CCS. In other words, there are episode records in the clinical and staging data file that have no linking records in the current clinical status file. This may occur if a practice did not report any data to the OCMR for a particular episode, or if the practice reported clinical and staging data, but did not report CCS data. As noted above, many beneficiaries have multiple CCS records, either per episode or across multiple episodes, because CMS required practices to report CCS at least once for each episode, as well as whenever there was a change in the CCS.

The following information is available in the current clinical status file:

- CCW beneficiary identification number (BENE_ID),
- ICD-10 cancer diagnosis code reported by participant (ICD-10_CNCR_DGNS_CD_RPTD),
- Initial diagnosis date (INITL_DGNS_DT),
- Current clinical status (CRNT_CLNCL_STUS), and
- Current clinical status date (CRNT_CLNCL_STUS_DT)

3.3 Linking OCM Data Files

The OCM data files contain identifier fields that researchers can use to join the files together. Researchers can link the clinical and staging data file and the current clinical status data file by the following variables: BENE_ID, ICD10_DX_REPORTED, and INITIAL_DIAGNOSIS_DATE.

4.0 Linking with Other CCW Data Files

By design, all beneficiaries initiating an OCM episode are Medicare FFS beneficiaries with Part A and B coverage. Many of them also had Medicare Part D coverage.

CCW adds a unique CCW beneficiary identifier (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 available data types, thus allowing data users to link the OCM data to beneficiary and claims data in the CCW.

The unique CCW beneficiary identifier field is specific to the CCW and does not apply to any other identification system or data sources. CCW encrypts this identifier and all data files before delivering the data files to researchers.

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

The CCW Medicare enrollment data file is the Master Beneficiary Summary File (MBSF) that uses the CMS Common Medicare Environment (CME) database as its source. The MBSF contains many enrollment and other person-level variables in file "segments." These segments are separate components of the file researchers may request. The data dictionary and codebook on the CCW website describe the variables contained in the MBSF.

The CCW team creates the MBSF for each calendar year. The MBSF contains demographic, entitlement, and enrollment data for beneficiaries who 1) CMS documents are alive for some of the reference year and 2) enrolled in the Medicare program during the file's reference year. Reference year refers specifically to the calendar year accounted for in the MBSF. So, for example, the 2020 MBSF covers the year 2020 — that is the reference year.

This essential information for most study denominators appears in the base A/B/C/D segment of the MBSF. For each of the MBSF file segments, there is one record for each BENE_ID. The additional segments of MBSF are 1) CCW Chronic Conditions, 2) CMS Other Chronic or Potentially Disabling Conditions (OTCC), 3) Cost and Use, and 4) National Death Index (NDI).[1]

Researchers may wish to obtain MBSF data fields for a population they identify within the OCM data files. Use the BENE_ID to perform this linkage. Remember that the OCM episodes included in these data files occurred in different time frames during July 2017–June 2022. Researchers should select the year(s) of the MBSF that coincide with the six months of each episode (e.g., using the EPISODE_BEG_DT and EPISODE_END_DT).

[1] Researchers may only use the NDI files within the CCW Virtual Research Data Center (VRDC).

4.2 Medicare Part A and B Claims

The CCW includes Medicare institutional and non-institutional claims, and Medicare Part D prescription drug fill events. CMS historically limited the Medicare claims found in the CCW to FFS Part A and B claims only. The Data Dictionaries tab on the CCW website describes the variables in the FFS claims files; researchers may also reference the *CCW Medicare Administrative Data User Guide* on the CCW website.

The CCW team adds key variables to the data files to help researchers join them together as appropriate (e.g., the unique CCW-assigned beneficiary identifier [BENE_ID], the claim identifier [CLM_ID], the claim line/record number [CLM_LINE_NUM]). The CCW team uses the last date on the claim, referred to as the CLM_THRU_DT, to partition the claims into calendar year files.

Researchers may wish to obtain FFS claims data for a population they identify within the OCM data files. If interested in claims for a beneficiary population, they should use the BENE_ID to perform this linkage. Remember that the OCM episodes included in these data files occurred in different time frames during July 2017–June 2022. Researchers may wish to examine claims data before, during or after the duration of OCM or the OCM episode. Researchers should select the months and year(s) of the claims files to correspond with the desired pre/post period using the record date within the file (e.g., the EPISODE_BEG_DT and EPISODE_END_DT in the clinical and staging data file).

5.0 Receiving CCW Data

This section describes the content and format of the CCW Medicare OCM data package that the CCW team will make available to researchers. The CCW team provides data files to the researcher in the following formats.

5.1 Within the CCW Virtual Research Data Center (VRDC)

The OCM files are available in the VRDC SAS library called CMMI_OCM. This library contains two data files:

OCM_CLINICAL_DATA.sas7bdat

OCM_ASSOCIATED_CCS.sas7bdat

Each data set includes the nine performance periods (third through eleventh) from July 2017–June 2022.

5.2 Physical Shipment of Data

Some researchers receive a physical data shipment from the CCW team. There are one or more folders on the physical media, each containing multiple files. CCW organizes the folders by request number as depicted below:

XXXXX (folder with your CCW data request number)
Extract file documentation

The researcher will find a folder named 2022 inside the request number folder that contains the OCM data. If the data request contains additional types of data besides OCM, there could be additional folders.

The CCW team creates password protected executable files (self-decrypting archives [SDA]) that contain the two OCM data files.

Inside the 2022 folder, there is a Read Me file and the OCM SDA (reference Table 2 and Table 4).

The naming convention for the SDA is as follows:

res<XXXXXXXXX>req<XXXXXX>_<YYYY>_<FTYPE>

Researcher DUA#
CCW request #
Year of data
File type

For example, if the DUA # was 000077777, the CCW request number was 012345, the year will always be 2022 for OCM and the file type is OCM.

The folders and data files would look like this:

12345
    2022
    READ_ME_FIRST_REQ12345_2022.txt
    res000077777_req012345_2022_OCM.exe

Table 2. Format and naming convention for the CCW files

File	File description
READ_ME_FIRST_REQ12345_2022.txt	This is a text file that describes the files contained in the output package. Filename example: READ_ME_FIRST_REQ12345_2022.txt
res000077777req012345_2022_OCM.exe	This is the SDA executable that researchers must run to decrypt and decompress the OCM data files. In this example, 000077777 is the DUA number, 012345 is the request number, and 2022 is the year of the data. This executable includes v8 SAS read-in program, the .psv file, and .fts file containing the layout and record counts.

Table 3. OCM SDA contents

File	File description
ocm_clinical_data_res<0000nnnnn>_req<0nnnnn>_2022.psv ocm_clinical_data_res<0000nnnnn>_req<0nnnnn>_2022.fts ocm_clinical_data_read_v8.sas	This set of files includes the OCM Clinical and Staging .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in program.
ocm_associated_ccs_res<0000nnnnn>_req<0nnnnn>_2022.psv ocm_associated_ccs_res<0000nnnnn>_req<0nnnnn>_2022.fts ocm_associated_ccs_read_v8.sas	This set of files includes the OCM Current Clinical Status .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in program.

In addition to the specific data files the researcher requested, the CCW team includes a decryption resource file in the deliverable package. Table 4 shows this file.

Table 4. CCW resources accompanying data files

File	Description
Decryption instructions.pdf	This document contains instructions for decrypting/uncompressing the data files.

The encryption technique for files extracted from the CCW uses Pretty Good Privacy (PGP) command line software. 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.

After the CCW team ships the data to the researcher, they send the password to decrypt the archive to the researcher via email. Each researcher request has a unique encryption. The CCW team never packages the password and the data media together. To decrypt the data files, the researcher accesses the email containing the decryption password. The data package contains detailed instructions for using this password.

6.0 Where to Get Assistance

Researchers interested in working with CCW data should contact ResDAC. 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. Visit the ResDAC website at http://www.resdac.org for additional information.

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

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

The CCW Help Desk staff 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 — List of Acronyms

Acronym	Definition
BENE ID	CCW beneficiary unique identification number
CCS	Current clinical status
CCW	Chronic Conditions Warehouse
CME	Common Medicare Environment
CMMI	Center for Medicare & Medicaid Innovation (CMS Innovation Center)
CMS	Centers for Medicare & Medicaid Services
DUA	Data Use Agreement
E&M	Evaluation and management
FFS	Fee-for-service
HCPCS	Healthcare Common Procedure Coding System
HMO	Health maintenance organization
MBI	Medicare beneficiary identifier
MBSF	Master Beneficiary Summary File
MEOS	Monthly enhanced oncology services
NDI	National Death Index
OCM	Oncology Care Model
OCMR	OCM data registry
OTCC	Other Chronic or Potentially Disabling Conditions
PBP	Performance-based payment
PFFS	Private fee-for-service
PGP	Pretty Good Privacy
PHE	Public health emergency
PPO	Preferred provider organizations
ResDAC	Research Data Assistance Center
RIF	Research Identifiable File
SDA	Self-decrypting archive
SNP	Special needs plans
TNM	Tumor-nodal disease-metastasis
VRDC	Virtual Research Data Center