Chronic Conditions Warehouse

Your source for national CMS Medicare and Medicaid research data

Chronic Conditions Warehouse

Million Hearts® Cardiovascular Disease (CVD) Risk Reduction Model Medicare Data Files User Guide

MAY 2023 │ VERSION 1.0

This page intentionally left blank.

Revision Log

Date	Changed by	Revisions	Version
May 2023	K. Schneider	Created initial document	1.0

Table of Contents

List of Tables

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). The Centers for Medicare & Medicaid Services (CMS) Innovation Center conducted a demonstration project from 2016–2021 called the Million Hearts Cardiovascular Disease (CVD) Risk Reduction Model. This project was a randomized controlled trial that sought to bridge a gap in cardiovascular care for Medicare beneficiaries by providing targeted incentives for health care practitioners to engage in beneficiary CVD risk calculation and population-level risk management. Instead of focusing on the individual components of risk, participating organizations engaged in risk stratification across a beneficiary panel to identify those at highest risk for atherosclerotic cardiovascular disease (ASCVD). There were 319 participating organizations involved in the Million Hearts CVD Risk Reduction Model.

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 obtained the Million Hearts Model data files from CMS. From this source data, CCW has prepared data files to disseminate to researchers and certain government agencies who CMS have approved under a Data Use Agreement (DUA) to obtain Medicare Million Hearts Model data for research purposes. The CCW Million Hearts CVD Risk Reduction Model Medicare data files contain identifiable information and 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 users with information to clarify their work with the CCW Medicare Million Hearts data files. Appendix A lists abbreviations used in this document.

Chapter 1. Background

The CMS Innovation Center conducted a demonstration project for the Million Hearts® CVD Risk Reduction Model. This demonstration project supported both the Million Hearts model's goal to prevent one million heart attacks and strokes and CMS' objective to identify and spread better models of care delivery and payment. This model aimed to improve quality while maintaining budget neutrality for Medicare beneficiaries ages 40–79 who had not had a previous heart attack or stroke.

The Million Hearts CVD Risk Reduction Model proposed an innovative way of lowering CVD risks across the population. Rather than paying health care practitioners to screen for blood pressure, cholesterol, or other risk factors individually, the Million Hearts CVD Risk Reduction Model generated individualized beneficiary risk scores and providers were to develop mitigation plans for eligible Medicare FFS beneficiaries. Participating organizations included primary care and cardiology practices, health centers, and hospital outpatient departments.

CMS received a total of 762 applications from provider organizations to participate in this five-year model. After a rigorous review and selection process, 516 organizations were chosen to participate. Randomization resulted in 260 organizations assigned to the intervention group and 256 organizations assigned to the control group. Selected participants were comprised of provider organizations in 47 states, Puerto Rico, and the District of Columbia. Among the 516 organizations, 345 (of the 516) participated in the first two model years. As of December 2020, about 60% of the 516 randomized organizations formally remained in the model, meaning they had not withdrawn, or CMS had not terminated them.

The model design included incentive payments to providers. Control practices received a one-time payment of $20 per beneficiary. The intent was for control group practices to report annually in years 1, 2, 3, and 5. For the intervention group, practices were to report every six months.[1] Providers in the intervention group received a per beneficiary per month (PBPM) payment that CMS determined according to the reduction in ASCVD risk compared to the baseline ($5 PBPM or $10 PBPM, etc.)

This project is no longer active. The Million Hearts demonstration project was conducted over a five-year period, beginning in January 2017, and ending in December 2021. CMS documents additional details related to the model design,[2] as well as annual evaluation reports pertaining to the model (reference https://innovation.cms.gov/data-and-reports/2022/mhcvdrrm-fourthannevalrpt, for example).

A. CVD Risk Scores

Each participating provider, whether part of the intervention or control group, used a modified version of the 10-year American College of Cardiology/American Heart Association (ACC/AHA) ASCVD pooled cohort risk calculator[3] to identify the CVD risk-score for all eligible Medicare FFS beneficiaries. For beneficiaries identified as high-risk, the provider was to engage in shared decision making, develop individual risk modification plans, and re-assess their risk on an annual basis throughout the five years of the model.

---

[1] CMS. "Million Hearts® Cardiovascular Disease Risk Reduction Model FAQs." https://innovation.cms.gov/files/x/mhcvdrrm-faqs.pdf

[2] CMS.gov Innovation Center. Million Hearts: Cardiovascular Disease Risk Reduction Model at https://innovation.cms.gov/innovation-models/million-hearts-cvdrrm

[3] An example of the current ASCVD risk estimator tool is at: https://tools.acc.org/ascvd-risk-estimator-plus/#!/calculate/estimate/

Control practices were not asked to implement the ACC/AHA ASCVD risk calculation or evidence-based risk modification but were asked to submit clinical data on all eligible Medicare FFS beneficiaries for comparison against intervention practices.

The ASCVD ten-year pooled cohort risk calculator used the following variables to calculate risk:

- Age
- Race
- Total and high-density lipoprotein (HDL) cholesterol levels
- Low-density lipoprotein (LDL) cholesterol
- Systolic blood pressure
- Use of statin therapy
- Antihypertensive medication
- Use of aspirin therapy
- Smoking status
- Diabetes status

Intervention and control organizations enrolled beneficiaries by collecting required demographic and clinical data needed to calculate the beneficiary's CVD risk. The beneficiary was considered enrolled as of the date of the visit that the organization collected the required clinical data elements to submit to the Million Hearts Data Registry (MHDR).

B. Million Hearts Data Collection/Registry

Participating providers submitted both the intervention and control practices data to the MHDR using the Million Hearts registry web interface. They provided beneficiary data by manual input by using Million Hearts beneficiary bulk upload templates for clinical and demographic data and uploading Consolidated Clinical Document Architecture (CCDA) files.

On a scheduled basis, the beneficiary data contained within the registry was copied and exported to the CCW environment, and adjusted beneficiary information was updated by the Million Hearts data contractor. This allowed the implementation and evaluation groups to perform required analysis on the exported registry data.

An enrolled beneficiary was a person for whom the organization uploaded data to the MHDR when the beneficiary had a baseline visit with the organization, as well as the demographic and clinical data needed to determine the beneficiary's baseline CVD risk. To be considered enrolled, CMS also validated each beneficiary's enrollment using claims data to confirm that the beneficiary 1) did indeed have a visit with a provider from the organization near the time listed and 2) met model eligibility criteria that could be replicated in enrollment and claims data. Medicare FFS beneficiaries met model eligibility criteria if they were ages 40 to 79, had no evidence of a prior heart attack or stroke, had Medicare as their primary payer, did not have ESRD, and were not receiving hospice benefits.

For the purposes of the Million Hearts CVD Risk Reduction Model, CMS defined "high-risk" beneficiaries as individuals with an ACC/AHA 10-year ASCVD risk score of greater than 30%. The ACC/AHA guidelines still held true that all individuals should have a risk score less than 7.5%. CMS was attempting to intervene on the critically high-risk group.

CMS expected that practices would provide the necessary care as clinically appropriate for beneficiaries with risk scores between 7.5–29%.[4]

Control organizations reported clinical indicators (but not a risk score) for visits with beneficiaries attributed to the model.

B.1 Key concepts

- CMS identifies providers as National Provider Identifiers (NPIs) within the practices (organizations) that participated in the Million Hearts model.
- Intervention group (organizations) — organizations (provider practices) that applied and CMS randomized to the study group (as opposed to usual care/control group). Not all organizations meaningfully participated; that is, not all submitted data to MHDR.
- Aligned (alignment) — CMS aligns beneficiaries with their provider practice; it informs which provider calculated the cardiovascular disease risk score for the beneficiary and subsequently that provider/clinician provided care to the beneficiary that resulted in a decrease (or increase, or no change) in the beneficiary risk score.
- Attributed (beneficiaries; population) — is CMS' best approximation of those eligible for the Million Hearts model, based on Medicare claims and enrollment data within a particular time period. CMS used Medicare claims to attribute beneficiaries to providers/organizations that participated in the model.
- Enrolled (beneficiaries) — enrolled means the organization (practice) reported the beneficiary to the MHDR and the CMS validated the beneficiary's enrollment record. To enroll a beneficiary, an organization had to upload data to the registry about when the beneficiary had a baseline visit with the organization and provide the demographic and clinical data needed to determine the beneficiary's baseline CVD risk. An organization that was part of the intervention group or control group enrolled the beneficiaries. Organizations may have enrolled beneficiaries in different performance periods (time frames) within the five-year study period.

The (beneficiary) Enrollment, Validation, Alignment, and Adjudication (EVAA) analysis was performed by the Million Hearts data contractor. The alignment file received from the registry contained beneficiary statuses determined by the Million Hearts registry (these statuses can be found in the REASON_CODE data field of the alignment file(s) and contain information on beneficiaries' program standing, such as lost-to-follow-up for beneficiaries who missed their follow-up visits). The updated beneficiary information — their eligibility, alignment, and reason codes — were updated in the respective ELIGIBLE, ALIGNED, and REASON_CODE field values of the alignment file. In addition, a list of beneficiaries who should have been enrolled into the Million Hearts program but were not (eligible but not enrolled beneficiaries) identified as part of EVAA were also appended to the alignment file.

B.2 Project Data files

The data from the Million Hearts model is a set of seven linkable data file types. The data includes detailed information about the beneficiaries, physician practices, specific providers, and clinical visits. The files do not contain names, beneficiary street addresses, or contact information.

Table 1 provides a list of data file types created as part of the Million Hearts model, as well as explanations and data field context in the following sections.

---

[4] CMS. "Million Hearts® Cardiovascular Disease Risk Reduction Model FAQs." https://innovation.cms.gov/files/x/mhcvdrrm-faqs.pdf

Table 1. Million Hearts project data file types

File type	File description
Clinical (raw)	The raw clinical data file contains all beneficiary clinical data uploaded by the participating practices throughout the program. It contains all data processed through the registry system.
Clinical (final)	The clinical file is the latest (final) data extract snapshot of beneficiary clinical data at the end of the program
Demographic	The demographic file contains all beneficiary demographic information entered through manual entry and uploaded by the participating practices
Visit	The visit file contains all data extracted to create a beneficiary visit and includes both uploaded data and manually entered data
Alignment	The alignment files contain the eligibility and alignment status of beneficiaries and a list of eligible-but-not-enrolled patients for the participating practices to report on. They contain an associated reason code for status changes. There is one final alignment file, as well as a file for each of the 10 EVAA cycles. Appendix B contains the list of status and reason codes identified for the Million Hearts model
Practice	The practice list file contains the list of all practices enrolled into the MHDR throughout the model's duration
NPI	The provider list file contains the list of all providers (identified by their NPIs) enrolled into the MHDR including the activation and deactivation dates.

About 30% of the high- and medium-risk beneficiaries in the intervention group were enrolled by organizations that either formally withdrew as of December 2020 or no longer reported to the registry in 2020.

After the model launched, the number of intervention organizations reporting visit data to the registry (a requirement to receive incentive payments) declined considerably over time, from 174 in the first six months of the model to 63 by the last six months of 2020.

Organizations that actively participated through 2020 tended to be larger, based on the average number of providers reported in the organizations' Million Hearts Model application, were less likely to be in a rural location, and more likely to be classified as a specialty or multispecialty practice.[5]

---

[5] Peterson, G. et al. "Evaluation of the Million Hearts® Cardiovascular Disease Risk Reduction Model: Fourth Annual Report." February 2022. https://innovation.cms.gov/data-and-reports/2022/mhcvdrrm-fourthannevalrpt

Chapter 2. CCW Million Hearts Data Files

CMS uses the CCW to develop and manage CMS research data resources. The CCW team creates Million Hearts CVD Risk Reduction Model Medicare Data files from the CMS Million Hearts model source data and disseminates it as a suite of 17 data files. The CCW team copies the Million Hearts source data to the CCW environment and assigns the CCW BENE_ID; the BENE_ID allows for linkage to other CCW data products (e.g., Medicare enrollment and claims). A small number of records in the demographic, clinical and visit files did not contain sufficient information from the Million Hearts project to result in a beneficiary match and therefore do not have a BENE_ID. All Million Hearts files with beneficiary information have the unique Million Hearts Model-Assigned Patient identifier (called the UNIQUE_BENEFICIARY_ID), even if a CCW BENE_ID is not available.

Throughout this Chapter, when the authors identify a particular data variable by name, the specific SAS name appearing in all capitals identifies it.

A. Practice Data

The practice file contains the list of all provider practices, or participating organizations, enrolled into the registry throughout the model's duration. Each practice (PRACTICE_ID and PRACTICE_NAME) has a GROUP_TYPE that identifies them as either participating in the intervention or control group.

B. Provider (NPI) Data

The provider file contains the list of all the providers, using the provider's National Provider Identifier (NPIs), enrolled into the registry including the activation and deactivation dates. The participating organization (PRACTICE_ID) enrolls the NPIs in the registry.

C. Demographic Data

The demographic file contains all beneficiary demographic information entered by the organization at the time the practice enrolled the patient in the Million Hearts registry. There is one record for each beneficiary. However, it is possible for the same beneficiary to appear in the file twice, if two or more organizations tried to enroll the same beneficiary. EVAA records are used to compare against data that providers submit. In some cases, the records may have different or more data included but they reference the same beneficiaries from the raw provider submitted data.

The UNIQUE_PATIENT_ID identifies each beneficiary enrolled in the Million Hearts program, and for records with sufficient information, the CCW files assigned the beneficiary a BENE_ID.

Not all beneficiaries in this file met the eligibility criteria for the Million Hearts model (i.e., presence of the beneficiary record in this file does not mean the model included the beneficiary in either the intervention or control group).

D. Clinical Data

There are two versions of the clinical file. The first is a raw data file that contains the data submitted by the practice; the second (or final) file is a processed version of the raw data. The layout for both files is the same. There is one record for each CVD risk factor entered by the practice for each beneficiary. In other words, a beneficiary has a record for each of the (up to 10) CVD risk factors that the practice assessed on a given date (ACTIVITY_DATE). The data identifies the NPI and practice (both PRACTICE_ID and PRACTICE_NAME).

- **Final clinical data** — CMS applied the Million Hearts business rules to the raw data and extracted and translated selected information into custom codes (field called CODE). Reference Table 2 contains a list of CODE values and their descriptions.

Table 2. Code values for final clinical file

CODE	Description
AFB.NO	Whether or not the patient has atrial fibrillation
AFB.YES	Whether or not the patient has atrial fibrillation
ASP.NO	Whether or not the patient is taking aspirin
ASP.YES	Whether or not the patient is taking aspirin
DBT.NO	Whether or not the patient has diabetes
DBT.YES	Whether or not the patient has diabetes
HDL	High-density lipoprotein
LDL	Low-density lipoprotein
SBP	Systolic blood pressure
SCM	Stress cardiomyopathy
SMK.NO	Smoking — whether or not the patient is a smoker
SMK.YES	Smoking — whether or not the patient is a smoker
STT.NO	Statin — whether the patient is on a Statin
STT.YES	Statin — whether the patient is on a Statin
TCH	Total cholesterol
THY.NO	Thyroid
THY.YES	Thyroid

The VALUE field contains the actual laboratory or blood pressure value observed for the CODE. For example, if the clinical record for the beneficiary has CODE = LDL, then the VALUE field is populated with the laboratory value for the LDL.

- **Raw clinical data** — this is the provider-submitted data and CMS has not edited or cleaned it. For most purposes, data users should use the final clinical data rather than this raw data file

E. Visit Data

The visit file contains information submitted by the practice for each beneficiary visit. This file includes the ASCVD calculated[6] CVD risk-score at the time of the visit (RISK_SCORE) as well as the baseline score (BASELINE_SCORE) and

---

[6] An example of the current ASCVD risk estimator tool is at: https://tools.acc.org/ascvd-risk-estimator-plus/#!/calculate/estimate/

the categorization for the baseline score into a risk grouping (BASELINE_GROUP) of high, medium or low.[7] If the visit is after the baseline visit, then the RISK_SCORE includes the value for that visit. Otherwise, at baseline, the two risk score fields match (i.e., BASELINE_SCORE equals RISK_SCORE).

The file is at the visit level, with one row for each visit for each beneficiary.

F. Alignment Data

The Million Hearts model employed a complex EVAA analytic process. During this process, the beneficiary data that providers submitted to the Million Hearts registry was validated against claims data. In some cases, the records may have different or more records ; from this validation step the beneficiary alignment (attribution) and program eligibility were determined. The EVAA process occurred semiannually, each October and April. There was a total of 10 EVAA analytical cycles, between October 2017 through April 2022 (two per year for the five years of the project). Table 3 lists the date when each of the 10 semi-annual alignment cycle data files was produced. There is also a "final" alignment file, which represents beneficiary alignment and eligibility at the end of the model.

The record layout is the same for each of the eleven alignment files. The alignment file contains the eligibility (variable called ELIGIBLE) and alignment status (ALIGNED)of beneficiaries and a list of eligible-but-not-enrolled patients for the participating practices to report on. The REASON_CODE contains one or more values to indicate the reason(s) for a patient status change (e.g., patient lost to follow-up). Appendix B presents the list of valid REASON_CODE values and their meaning.

Depending on the study objectives, researchers may or may not need to identify the exact beneficiaries who were verified to be part of the provider organization cohort during any particular time period (alignment period).

Table 3. Alignment file cycle

EVAA cycle	Date Million Hearts registry generated the file
#1	01 Dec 2017
#2	01 May 2018
#3	01 Nov 2018
#4	01 May 2019
#5	01 Nov 2019
#6	01 May 2020
#7	03 Nov 2020
#8	01 May 2021
#9	01 Nov 2021
#10	01 May 2022

---

[7] CMS FAQ states, "For the purposes of the Million Hearts® Cardiovascular Disease Risk Reduction Model, CMS is defining "high-risk" beneficiaries as individuals with an ACC/AHA 10-year ASCVD risk score of greater than 30%." Reference https://innovation.cms.gov/files/x/mhcvdrrm-faqs.pdf

G. Linking Million Hearts Data Files

Each of the Million Hearts data files contain identifier fields that researchers can use to join the files together. PRACTICE_ID and corresponding PRACTICE_NAME identifies each of the practices, and these two variables are included in each of the seven file types. Each of the beneficiaries has a UNIQUE_PATIENT_ID, and the system has assigned nearly all records the CCW BENE_ID.

Chapter 3. Linking with Other CCW Data Files

By design, all beneficiaries enrolled in the Million Hearts Model are Medicare FFS beneficiaries with Part A and B coverage. About 70% also had Medicare Part D coverage.[8]

CCW removes the Medicare beneficiary health insurance claim (HIC)[9] 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 Million Hearts 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 source. CCW encrypts this identifier and all data files before delivering the data files to researchers.

A. 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.[10] The MBSF contains many enrollment and other person-level variables contained in file "segments." These segments are separate components of the file researchers may request. The data dictionaries on the CCW website (reference https://www2.ccwdata.org/web/guest/home/) 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 part 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).[11]

Researchers may wish to obtain MBSF data fields for a population they identify within the Million Hearts' data files. Use the BENE_ID to perform this linkage. Remember that Million Hearts' organizations (practices) may have enrolled beneficiaries at a different time frames during the five-year study period between 2017–2022. Researchers should select the year of the MBSF to match the record date within the file (e.g., the CREATE_DATE in the demographic file or the VISIT_DATE in the visits file).

---

[8] Peterson, G. et al., "Evaluation of the Million Hearts® Cardiovascular Disease Risk Reduction Model: Fourth Annual Report." February 2022. https://innovation.cms.gov/data-and-reports/2022/mhcvdrrm-fourthannevalrpt

[9] CMS began using a new Medicare beneficiary identifier (MBI) in place of the HIC, starting in 2018.

[10] A CCW White Paper, *Medicare Enrollment: Impact of Conversion from EDB to CME* contains a description of the rationale and impact of the data conversion from the CMS Enrollment Database (EDB) to the CME for producing the MBSF.

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

B. Medicare Part A and B Claims

The CCW system includes Medicare institutional and non-institutional claims, as well as 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 contains a description of the variables in the FFS claims files; users may also reference the *CCW Medicare Administrative Data User Guide* on the CCW website.

MA (Part C) encounter data RIFs are available to researchers starting with 2015. Medicare Advantage Organizations (MAOs) are private managed care plans, such as health maintenance organizations (HMOs), preferred provider organizations (PPOs), private fee-for-service plans (PFFS), and special needs plans (SNPs) that provide Medicare Part A and Part B services. MAOs submit data to CMS that the CCW team uses to create the RIFs. Reference the Medicare Encounter Records data dictionaries along with a *CCW Medicare Encounter Data User Guide* on the CCW website.

The CCW team adds key variables in the data files to help researchers join them together as appropriate (e.g., the unique CCW-assigned beneficiary identifier [variable called 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 claims data for a population they identify within the MH Data files. If interested in claims for a beneficiary population, they should use the BENE_ID to perform this linkage; if interested in a provider population they should use the TIN(s) or NPI(s) from the MH data to join to the PRVDR_NUM or any of the NPI fields in the CCW claims files. Remember that MH organizations may have enrolled beneficiaries at a different time frames during the five-year study period between 2017–2022. Researchers may wish to examine claims data before, during or after the MH project. Researcher 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 VISIT_DATE in the visits file or the ACTIVITY_DATE in the final clinical file).

Chapter 4. Receiving CCW Data

This section describes the content and format of the CCW Medicare Million Hearts data package (the CCW data physically delivered to researchers). The CCW team delivers files to the researcher in the following format. Table 4 lists descriptions of each of these contents. The CCW team delivers all 17 files in one data package. Although files encompass the Million Hearts project from 2017–2021, the file naming convention only has the 2021 year.

A. Format

The CCW team delivers files to the researcher in the following format. There are one or more folders, each of which contains multiple files. The folders are:

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

Researchers have additional folders if they requested additional types of data besides the Million Hearts' data package.

The CCW team inserts all the Million Hearts' data files within the folder with the CCW data request number (reference Table 4 and Table 5). The year of the Million Hearts' data package is 2021; if you also obtain other Medicare data files, there is a separate sub-folder for each year of enrollment or claims data requested.

Executable files (self-decrypting archive [SDA]) hold all the data files. Researchers need to enter a password to extract the files.

The naming convention for the SDA is as follows:

> Res<XXXXXXXXXX>req<XXXXX>_<YYYY>_<FTYPE>
>
> Researcher DUA#   Year of data
>
>   CCW request #   File type
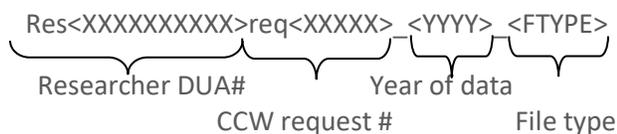
For example, if the DUA # was 0000077777, the CCW request number was 12345, and the researcher obtained the Million Hearts Model (MHM) data files, the folders and data files would look like this:

📁 12345
    📁 2021
        READ_ME_FIRST_REQ12345_2021.txt
        res000077777req012345_2021_MHM

The naming convention uses abbreviations to indicate the types of data files (enrollment, claims, or other data files).

Table 4. Format and naming convention for the CCW files

File	File description
READ_ME_FIRST_REQ12345_2021.txt	This is a text file that describes the files contained in the output package. Filename example: READ_ME_FIRST_REQ12345_2021.txt
res000077777req0121345_2021_MHM.exe	This is the SDA executable that researchers must run to decrypt and uncompress the MHM data files. In this example, 000077777 is the DUA number, 012345 is the request number, and 2021 is the year of the data. This executable includes v8 SAS read-in programs, the .psv file, and .fts file containing the layout and record counts.

MHM files

File	File description
mhm_align_res<0000nnnnn>_req<0nnnnn>_2021.psv mhm_align_res<0000nnnnn>_req<0nnnnn>_2021.fts mhm_align_read_v8.sas	This set of files includes the final alignment .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
mhm_align_{01-10}_res<0000nnnnn>_req<0nnnnn>_2021.psv mhm_align_{01-10}_res<0000nnnnn>_req<0nnnnn>_2021.fts mhm_align_01_read_v8.sas, etc.	This set of 10 files includes the cyclical EVAA alignment files {for alignment periods 01-10; reference Table 3 for description} .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
mhm_clinical_res<0000nnnnn>_req<0nnnnn>_2021.psv mhm_clinical_res<0000nnnnn>_req<0nnnnn>_2021.fts mhm_clinical_read_v8.sas	This set of files includes the final clinical .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
mhm_clinical_raw_res<0000nnnnn>_req<0nnnnn>_2021.psv mhm_clinical_raw_res<0000nnnnn>_req<0nnnnn>_2021.fts mhm_clinical_raw_read_v8.sas	This set of files includes the raw clinical .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
mhm_demog_res<0000nnnnn>_req<0nnnnn>_2021.psv mhm_demog_res<0000nnnnn>_req<0nnnnn>_2021.fts mhm_demog_read_v8.sas	This set of files includes the demographic .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
mhm_npi_res<0000nnnnn>_req<0nnnnn>_2021.psv mhm_npi_res<0000nnnnn>_req<0nnnnn>_2021.fts mhm_npi_read_v8.sas	This set of files includes the NPI .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
mhm_practice_res<0000nnnnn>_req<0nnnnn>_2021.psv mhm_practice_res<0000nnnnn>_req<0nnnnn>_2021.fts mhm_practice_read_v8.sas	This set of files includes the practice .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.
mhm_visit_res<0000nnnnn>_req<0nnnnn>_2021.psv mhm_visit_res<0000nnnnn>_req<0nnnnn>_2021.fts mhm_visit_read_v8.sas	This set of files includes the visit .psv (data) file, .fts (layout and record counts) file, and version 8 SAS read-in programs.

In addition to the specific data files the researcher requested, the CCW team includes a variety of resource files in the deliverable package. Table 5 shows these files.

Table 5. CCW resources accompanying data files

File	Description
Decryption instructions.pdf	This document contains instructions for decrypting/uncompressing the data files.
Tips on getting started with data.pdf	This document contains tips for using the CCW data.

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

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.

Chapter 5. 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 not able to answer questions, the technical 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 the researcher 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
ASCVD	Atherosclerotic cardiovascular disease
CCW	Chronic Conditions Warehouse
CME	Common Medicare Environment database
CMS	Centers for Medicare & Medicaid Services
CVD	Cardiovascular disease
DUA	Data Use Agreement
ESRD	End-stage renal disease
EVAA	Enrollment, Validation, Alignment, and Adjudication (analytic process)
FFS	Fee-for-service (claims)
HIC	Health insurance claim number
MA	Medicare Advantage
MAO	Medicare Advantage Organizations
MBI	Medicare beneficiary identifier
MBSF	Master Beneficiary Summary File
MH	Million Hearts®
MHM	Million Hearts Model (demonstration project)
MHDR	Million Hearts Data Registry
NPI	National Provider Number
PDE	Part D Prescription Events file
PII	Personally identifiable information
PO	Participating organization
ResDAC	Research Data Assistance Center
SDA	Self-decrypting archive

Appendix B — Code Sets

Valid values for REASON_CODE (within the alignment files). Codes that begin with "S" are status codes, those that begin with "SC" are supporting codes or subcodes.

Registry status	Status display in registry (patient grid)	Supporting codes (sub-codes)	Code definition
S001	Incomplete record	SC001	The temporarily not enrolled for cause beneficiary has now been accepted
		SC002	The pending acceptance beneficiary is accepted, but data is not complete
S002	Released for validation	SC003	PO updated HICN for verify HICN action
		SC023	PO updated record for verify beneficiary action
		SC022	PO updated MBI for verify MBI action
		SC033	Beneficiary lost to follow-up
		SC034	PO updated records for dropped action
		SC036	PO updated visit for verify visit action
S003	Enrolled	SC004	The pending acceptance beneficiary was accepted and completed
		SC005	The temporarily not enrolled for cause beneficiary has now been accepted
		SC006	Beneficiary status changed to aligned
	Low/Medium	SC015	The alignment record indicates that the low or medium risk patient is not aligned but eligible
S004	Pending acceptance	SC013	The beneficiary is pending acceptance
S005	Not aligned	SC014	Beneficiary status changed to not aligned
S006	Verify HICN	SC017	Verify HICN
S007	Ineligible	SC026	Exclusion: The beneficiary is not enrolled in Medicare FFS Parts A and B
		SC027	Exclusion: The beneficiary is not age 40–79 at time of enrollment
		SC028	Exclusion: The beneficiary is enrolled in the hospice benefit
		SC029	Exclusion: The beneficiary is enrolled in Medicare Advantage
		SC030	Exclusion: Medicare is not a primary payer
S008	Permanently ineligible	SC025	Exclusion: The beneficiary is not alive
		SC031	Exclusion: The beneficiary has ESRD
		SC032	Exclusion: The beneficiary has a previous heart attack or stroke
S009	Permanent: Not enrolled for cause	SC007	Significant comorbidities with a limited life expectancy of less than two years
		SC008	Unstable angina
		SC009	Class IV heart failure
		SC010	Peripheral vascular disease
		SC011	Prior heart attack or stroke

Registry status	Status display in registry (patient grid)	Supporting codes (sub-codes)	Code definition
S010	Temporary: Not enrolled for cause	SC012	One-time encounter/Will not return for follow-up
		SC013	Beneficiary declines to participate in the model
	Unavailable cholesterol	SC016	Cholesterol not available or out of range
S011	Lost to follow-up		Enrolled beneficiary lost to follow-up
S012	Dropped		Low/Medium beneficiary lost to follow-up
S013	Pending validation	SC033	Beneficiary was lost to follow-up
		SC023	PO updated record for verify beneficiary action
		SC034	PO updated records for dropped action
S014	Verify beneficiary	SC018	The DOB needs to be verified
		SC019	The gender needs to be verified
		SC020	The race needs to be verified
S015	Verify MBI	SC022	The MBI needs to be verified
S016	Verify visit	SC021	The visit date needs to be verified
		SC035	The NPI needs to be verified