CMS Research Identifiable File Data Use Agreement Policy Guide

Contents

1. Purpose

The purpose of the Centers for Medicare & Medicaid Services (CMS) Research Identifiable File (RIF) Policy Guide is to outline the policies that govern the disclosure of RIF data for research purposes. Researchers that request RIF data must also comply with the terms of the CMS Data Use Agreement (DUA).

2. Background

CMS makes certain data files available to internal and external stakeholders as allowed by federal laws, regulations, and CMS policy.

CMS maintains three different categories of data files: identifiable data files, limited data set files, and public use files. The privacy level of the data file dictates if a DUA is needed, the request process, and the level of review required:

1. Identifiable Data Files (IDFs) — IDFs contain protected health information (PHI) and/or personally identifiable information (PII) and are only available to certain stakeholders. IDFs can be physically shipped to a requester or accessed virtually. Research Identifiable Files (RIFs) are a type of IDF that are exclusively disclosed for research purposes in compliance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy Rule requirements at 45 C.F.R. § 164.512(i) and other applicable laws. IDF requests generally

require a DUA with CMS.

2. Limited Data Set (LDS) Files — LDS files contain PHI, but do not contain specific direct identifiers as set forth in the HIPAA Privacy Rule at 45 C.F.R. § 164.514(e)(2). CMS LDS files are only available for research purposes and require a DUA with CMS.
3. Public Use Files (PUFs) — PUFs (also called non-identifiable data files) do not contain PHI, PII, and are fully de-identified in accordance with HIPAA Privacy Rule requirements and the CMS cell suppression policy. PUF requests do not require a DUA and are not tracked by CMS. PUFs can be accessed freely on CMS' open data websites such as data.cms.gov and data.medicaid.gov.

3. Change Log

Change	Date
Policy Creation	February 12, 2024
Policy Update to Section 5.1.8	July 1, 2024

4. Definitions

Collaborating Organization: A Collaborating Organization works with the Requesting Organization, is involved in the research, and will be viewing or accessing unaggregated data or results that do not meet the CMS cell size suppression policy.

Data Custodian: Individual who will be responsible for ensuring that the environment in which the CMS data is stored complies with all applicable CMS data security requirements, including the establishment and maintenance of security arrangements to prevent unauthorized use. CMS requires only one data custodian per DUA.

Innovator: A researcher associated with a for-profit organization, or anyone conducting research with the intent to create a product or tool to be sold.

Personally Identifiable Information (PII): Information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual, in accordance with OMB Memorandum M-17-12 Preparing for and Responding to a Breach of Personally Identifiable Information (January 3, 2017). As noted in OMB M-17-12, the term PII is necessarily broad because there are many different types of information that can be used to distinguish or trace an individual's identity.

Protected Health Information (PHI): "PHI" or "Protected Health Information" is information, including demographic data, that relates to: the individual's past, present or future physical or mental health or condition, the provision of health care to the individual, or the past, present, or future payment for the provision of health care to the individual, and that identifies the individual or for which there is a reasonable basis to believe it can be used to identify the individual. PHI includes many common identifiers (e.g., name, address, birth date, Social Security Number). The HIPAA definition of PHI can be found at 45 C.F.R. § 160.103.

Requester: Individual authorized to sign agreements on behalf of the Requesting Organization. This person is often referred to as the 'legal signatory' and accepts all terms and conditions in

the DUA and attests that all the information contained in the request packet is accurate.

Requesting Organization: The organization with which the RIF DUA is established.

5. Research Identifiable File (RIF) Request Policies

CMS permits external stakeholders (referred to herein as "researcher(s)") to request access to non-public CMS data in the form of RIFs for research purposes, provided the researchers comply with the policies set forth in this Section 4.

5.1 Request Process

Researchers interested in requesting RIFs must work with the Research Data Assistance Center (ResDAC), the CMS technical assistance contractor, to complete the required research request packet. Researchers can find the required forms for the research request packet via the ResDAC Forms Generator Tool: https://resdac.org/request-forms-generator.

5.1.1 Submit Research Request Packet

Once a researcher finalizes a research request packet, ResDAC submits the packet on the researcher's behalf to CMS. All new requests are reviewed by the CMS Privacy Board.

5.1.2 Undergo CMS Privacy Board Review

The CMS Privacy Board reviews all potential disclosures of CMS RIF data for research purposes to ensure compliance with the requirements of the HIPAA Privacy Rule, the Privacy Act of 1974, other applicable federal laws and regulations, and CMS data policies. The CMS Privacy Board does not issue HIPAA waivers of authorization for external researchers. Researchers are required to provide documentation of a HIPAA waiver of authorization that meets the requirements of 45 C.F.R § 164.512(i)(1)(i) with the research request materials.

5.1.3 Pay CMS Fees

CMS charges fees to researchers to recoup the cost of making data available for research purposes. These fees are based on CMS's costs and allow the agency to continue to make data available to researchers. CMS's fees are standardized, and the agency is not able to offer discounts or concessions. CMS is only able to accept payment of fees through pay.gov.

RIF data fees are posted at: https://resdac.org/cms-fee-information-research-identifiable-data.

5.1.4 Receive Data Access

Researchers only receive access to CMS RIF data once their research request packet is approved by CMS, the RIF DUA is signed by all parties, and CMS fees are paid. The RIF DUA is the legal agreement between CMS and the researcher that documents the terms and conditions under which the CMS data may be used.

5.1.5 Publish Research

Researchers are required to publicly disseminate their research findings in compliance with the RIF DUA. CMS does not specify the mechanism for publication; researchers may publish in peer-reviewed journals or other public forums (e.g., blog, white paper).

5.1.6 Extend DUA

CMS RIF DUAs are only approved for a period of one year. The expiration date is included in the RIF DUA. Researchers may request to extend their DUA by submitting a RIF DUA Extension request via email. RIF DUA extension requests can only be submitted within 60

calendar days of the RIF DUA expiration date.

5.1.7 Amend DUA

After the research request packet is approved by CMS, researchers must notify CMS of any updates or changes to the RIF DUA or project information by submitting an amendment. Researchers can find the required paperwork for the different types of amendments at: https://resdac.org/managing-your-project-after-obtaining-data. The following amendments are reviewed by the CMS Privacy Board: (1) changing the description of the study (e.g., changes to the study cohort or aims) and (2) adding new data files that weren't previously requested for the study.

5.1.8 Close DUA

Researchers are required to close their RIF DUA by the RIF DUA expiration date or at the completion of their research project, if that occurs prior to the RIF DUA expiration date. DUAs that are not closed or extended within 30 days of the expiration date are in violation of the terms of the DUA and will be forced to close and the data destroyed. Requesting Organizations with an expired DUA will be unauthorized and will not be permitted to update or extend current DUAs or enter into a new DUA with CMS until the violation is resolved.

5.2 Available Data Files and Variables

While CMS collects and generates a significant amount of data in running our programs, only select data files and variables are available for research purposes. CMS has developed a standard set of data files with defined layouts and time periods that researchers may request. CMS generally does not develop custom files, file layouts, or data time periods for researchers. In addition, researchers cannot request future years of available data files.

The list of standard RIF data files is posted at: https://resdac.org/file-availability.

5.3 Linking to Other Data Sources

CMS must approve all data linkages between CMS RIF data and other data sets. Researchers may request linkages with a description of the linkage methodology to other data as part of the researcher's request packet to access CMS data.

Requests to link other beneficiary-identifiable data sources with CMS RIF data will only be approved if CMS performs the match. Researchers must submit identifiers from the other data source to CMS as part of a finder file. The finder file must only include the minimum necessary data elements about individuals to accomplish the purpose of the research initiative. Researchers must provide justification for the finder file and the PHI/PII needed to perform data linkages in advance of delivery of the finder file. All data submitted CMS, as well as linked data provided in response, for this purpose must comply with the Privacy Act of 1974, as amended, (5 U.S.C. § 552a), the HIPAA regulations (45 C.F.R. parts 160 and 164), and other applicable laws.

Researchers can submit one primary identifier or at least three secondary identifiers to CMS. Examples of primary identifiers include Social Security Number (SSN), Health Insurance Claim Number (HICN) and/or Medicare Beneficiary Identifier (MBI), and Medicaid Statistical Information System (MSIS) ID. Secondary identifiers include last name, gender, date of birth, zip code, and partial SSN/HICN. CMS will provide a crosswalk to link exact matches between the two sources. Please see the CCW Submission of Medicare Data Finder Files policy for additional information on requirements to submit a finder file to CMS to link data.

5.4 Data Access

CMS generally recommends that researchers access RIFs via the CMS Chronic Conditions Warehouse Virtual Research Data Center (CCW VRDC). The CCW VRDC is a secure CMS environment where researchers can access and analyze data. Currently, researchers also have the option to request physical copies of CMS data that will be stored and analyzed at the researcher's organization.

5.4.1 Physical Data Access Restrictions

The following RIFs and RIF cohort sizes are not eligible for physical data access:

- Master Beneficiary Summary File (MBSF): National Death Index (NDI) Segment;
- Medicaid Enrollee Supplemental File (MESF): National Death Index (NDI) Segment;
- Medicare Fee-for-Service (FFS) or Medicare Advantage Carrier Claims/Encounters for a cohort greater than 20% of the Medicare population;
- Medicare Part D Events for a cohort greater than 40% of the Medicare population; and
- Monthly COVID-19 research data files (Medicare FFS claims (Inpatient, Outpatient, Carrier, Durable Medical Equipment, Skilled Nursing Facility, Home Health Agency, Hospice), Part D Event File, Long Term Care Minimum Data Set 3.0, Master Beneficiary Summary File Base).

5.4.2 Physical Data Access Requirements

To request physical copies of CMS RIF data, a researcher must have an approved Data Management Plan Self-Attestation Questionnaire (DMP-SAQ). A copy of the DMP-SAQ approval letter is submitted with the research request packet. The DMP-SAQ documents the researcher's security and privacy controls to protect the privacy and security of physical copies of CMS data. For more information, please see: https://resdac.org/request-form/dmp-saq.

All physical copies of CMS data held by research organizations must adhere to all RIF DUA requirements. In addition, physical copies of CMS RIF data must be destroyed upon closure of the DUA, unless the data is also covered under another active DUA. Research organizations must attest to the destruction of the data by submitting a completed Certificate of Disposition (COD) to CMS.

5.4.3 Physical Data Access Reuse Restrictions

Researchers are permitted to "reuse" physical copies of CMS data held by their organization for a new research request or via an amendment to an existing research request if approved by CMS. Reuse requests must comply with the RIF DUA requirements and all policies described herein.

5.4.4 CCW VRDC Access Restrictions

Researchers are only permitted to access the data listed on their RIF DUA in the CCW VRDC.

Researchers may download aggregate, statistical information that is de-identified under the HIPAA Privacy Rule as described at 45 C.F.R. § 164.514(b) and that adheres to the CMS policy for cell size suppression from the CCW VRDC. All requests to download data from the CCW VRDC go through an output review to ensure the output complies with the terms of the researcher's RIF DUA.

5.4.5 Restrictions for Researchers Outside the United States

Researchers accessing CMS RIF data must be physically located in the United States (50 states, Washington, D.C., and the territories). Requesting Organizations receiving physical data files must store the data in a system based in the United States.

6. Special RIF Request Programs

6.1 Collaboration

Collaborating Organizations are permitted to collaborate on a research study with the Requesting Organization. Collaborating Organizations must be listed on the RIF DUA and the Collaborating Organization Supplement must be completed for each Collaborating Organization as part of the research request packet.

Researching Organizations requesting physical data are currently permitted to request that a second copy of the data be sent to the Collaborating Organization as part of the research request packet under the following circumstances: (1) the Collaborating Organization is non-profit; (2) the Collaborating Organization has an approved DMP-SAQ with CMS; and (3) the Requesting Organization pays the associated fees.

Collaborating Organizations must be actively involved in the research study.

6.2 Innovator/For-Profit Organization

The CMS Innovator/for-profit program allows for-profit organizations and organizations collaborating with for-profit organizations to use RIF data for research. This program also allows researchers to use RIF data to: (1) create a tool or product that will be sold, and/or (2) conduct analyses to support their business needs. The research study must be the primary basis for the request and all RIF data requested must be justified solely on the study protocol. Any product, tool, or business-related analyses must be secondary to the research.

An organization requesting to participate in the Innovator/for-profit program must complete the CMS Innovator Program Supplement as part of its research request packet. The CMS Privacy Board will review the request packet. Then, the CMS Chief Data Officer (CDO) conducts a second-level review after the CMS Privacy Board. The CDO's review focuses on whether the proposed use of the RIFs has the potential to exploit beneficiaries or create fraud and/or abuse in CMS programs.

Innovators are not permitted to access physical copies of CMS RIF data; they must use the CCW VRDC.

6.3 State Agencies

The CMS state agency program allows a state agency to request RIFs to fulfill its research purposes for a broader range of activities and programs. Specifically, the requesting state agency is permitted to reuse RIFs for additional research studies without CMS's direct approval. The requesting state agency may also further disseminate RIFs to other state agencies or entities conducting research that is directed and funded by the state. Under this program, a state agency can only request data for beneficiaries residing in the requesting state as well as a 5% national sample for benchmarking.

State agencies complete the State Agency Supplement as part of their research request packet. In this supplement, states decide whether to opt-in to the state agency program. State agencies requesting RIF data are not required to participate in the state agency program.

When states choose to opt-in to data sharing, the RIF DUA is modified via the State Agency Supplement, to permit the agency to reuse and further disseminate CMS data for additional research. State agencies who opt-in to data sharing are required to submit a quarterly log of state research using CMS data.

6.4 Federal Agencies

Federal agencies are permitted to request certain exceptions to CMS RIF policies as described in this section. Exceptions must be approved by CMS.

6.4.1 Available Data Files and Variables

Federal agencies may request CMS data beyond the standard RIFs. This includes data from other sources within CMS, non-standard data elements, and more timely data.

To request non-standard data, a federal agency must submit a letter of justification to the CMS CDO as part of its research request packet. CMS will review the request and determine if the request can be accommodated.

6.4.2 CMS Fees

Federal agencies are subject to CMS data fees. Fees may be paid via an Interagency Agreement (IAA) or through a contractor-relationship to pay fees via pay.gov. Federal agencies that are only reusing data from physical files are not charged fees.

Federal agencies that receive non-standard files may be subject to additional fees to cover the cost of making the data available.

Federal agencies accessing RIFs via the CCW VRDC are required to pay data fees.

6.4.3 For-Profit Contractors

CMS generally requires for-profit entities to access CMS RIF data via the CCW VRDC; however, federal agencies may collaborate with a for-profit contractor for data analysis support or as its data custodian and receive physical copies of the data.

6.4.4 CCW VRDC Access Restrictions

Federal agencies may request a waiver of output review in the CCW VRDC. To request an output review waiver, a federal agency must: (1) submit a letter of justification to the CMS CDO as part of its research request packet and (2) complete a DMP-SAQ to show that the data storage environment where the output will reside meets CMS privacy and security policies. All requests for a waiver of output review must have sufficient justification specifying why the removal of data from the CCW VRDC is essential for the project.

6.4.5 Disclosure of Linked Data

Federal agencies may request to further disseminate RIFs linked to other data (i.e., CMS RIF data linked to non-CMS data) for research purposes. A federal agency may also disseminate a limited sample of CMS data (e.g., a 5% sample) when disseminated with linked data files for research purposes.

To request permission to disseminate linked data files, a federal agency must:

- Establish a RIF DUA with CMS for the initial data disclosure; and
- Complete a Federal Agency Data Sharing Supplement as part of the research request packet.

Federal agencies may also request that CMS make a linked dataset created by the federal agency available through the CMS research request process. To request CMS dissemination of linked data files, a federal agency must submit a letter of justification to the CMS CDO as part of its research request packet. CMS will review the request and determine if the request can be accommodated. In certain cases, CMS may require funding from the agency to support CMS's activities.

6.5 Qualified Entities

Organizations that are approved as Qualified Entities (QEs) and have received CMS data under the CMS Qualified Entity program (also known as the Medicare Data Sharing for Performance Measurement Program) are permitted to request to "reuse" that data for research purposes. In general, QEs must comply with the RIF DUA as well as all policies described herein with respect to those reuses. However, QEs that are for-profit organizations or are collaborating with for-profit organizations are permitted to request to reuse physical copies of RIF data received through the QE program rather than access data through the innovator/for-profit program as described in Section 5.2 above.

6.6 CMS Contractors

CMS contractors occasionally receive physical copies of CMS data for the purposes of executing their contracts with CMS. Currently, CMS contractors are permitted to request to "reuse" that data for research purposes. Contractors must comply with the RIF DUA and all policies described herein with respect to those reuses. Contractors may only reuse data that is available in standard RIFs (see Section 4.2). Contractors may not use their access to CMS systems to create physical copies of CMS data for research if that data is not required for their CMS contract.

6.7 CMS Program Participants

Certain participants in CMS programs and initiatives (e.g., participants in Center for Medicare and Medicaid Innovation models) may receive physical copies of CMS data for the purposes of the CMS program or initiative. These participants are permitted to request to "reuse" that data for research purposes. These participants must comply with the RIF DUA and all policies described herein with respect to those reuses. These participants may only reuse data that is available in standard RIFs (see Section 4.2). These participants may not use their access to CMS systems to create physical copies of CMS data for research if that data is not required for the CMS program or initiative.