Medicare Program Integrity Manual
Chapter 2 – Data Analysis

Table of Contents
(Rev. 12772; Issued: 08-09-24)

<u>**Transmittals for Chapter 2**</u>

2.1 – Identifying Potential Errors - Introduction
(Rev. 10365; Issued: 10-02-20; Effective: 08-27-20; Implementation: 08-27-20)

A.   Contractors To Which This Section Applies

This section applies to MACs, UPICs, Recovery Auditors, and Supplemental Medical Review Contractor (SMRC).

B.   General

This chapter specifies resources and procedures to the MACs, UPICs, Recovery Auditors, and the SMRC. The contractors shall use these instructions to identify and verify potential errors to produce the greatest protection to the Medicare program. Contractors should objectively use analytical methodologies to evaluate potential errors and not take administrative action unless they have verified the error and determined that the error is a high enough priority to justify the action. They should also archive the error including supporting rationale for selection. (See Reliable Information in Pub. 100-08, Exhibits, Exhibit 4.)

C.   Review of Data

Data analysis is an essential first step in determining whether patterns of claims submissions and payments indicate potential problems. Such data analysis should include identification of statistical outliers in billing patterns within a well-defined group, or more sophisticated detection of patterns within claims or groups of claims that might suggest improper billing or payment.

Data analysis shall be undertaken as part of general surveillance and review of submitted claims, or shall be conducted in response to information about specific problems stemming from complaints, provider or beneficiary input, fraud alerts, reports from CMS, other MACs, or independent government and nongovernmental agencies.

2.2 – Data Analysis
(Rev. 10365; Issued: 10-02-20; Effective: 08-27-20; Implementation: 08-27-20)

A.   Contractors To Which This Section Applies

This section applies to MACs, UPICs, SMRC. This section does not apply to the Recovery Auditors. Recovery Auditors should follow the data analysis instructions listed in their Statement of Work.

B.   General

Data analysis is a tool for identifying actual or potential claim payment errors. Data analysis applies well-established statistical methods to claim information and other related data to identify potential errors and potential fraud by claim characteristics (e.g.,

diagnoses, procedures, providers, or beneficiaries) individually or at an aggregate level. Data analysis is an integrated, on-going component of MR and benefit integrity (BI) activity.

The MACs and UPICs ability to make use of available data and apply innovative analytical methodologies is critical to the success of the MR and BI programs. They should use research and experience in the field to develop new approaches and techniques of data analysis. The MACs and UPICs should have ongoing communication with other government organizations (e.g., QIOs and the State Medicaid agencies) concerning new methods and techniques.

Analysis of data should:

- Identify those areas of potential errors (e.g., services which may be non-covered or not correctly coded) that pose the greatest risk;

- Establish baseline data to enable the recognition of unusual trends, changes in utilization over time, or schemes to inappropriately maximize reimbursement;

- Identify where there is a need for an LCD;

- Identify where there is a need for targeted education efforts;

- Suggest claim review strategies that efficiently prevent or address potential errors (e.g., prepayment edit specifications or parameters);

- Produce innovative views of utilization or billing patterns that illuminate potential errors;

- Identify high volume or high cost services that are being widely overutilized. This is important because these services do not appear as an outlier and may be overlooked when, in fact, they pose the greatest financial risk;

- Identify program areas and specific providers for possible fraud investigations; and

- Determine if major findings identified by Recovery Auditors, CERT, and CMS represent significant problem areas in the MAC's jurisdiction.

This data analysis program shall involve an analysis of national data furnished by CMS as well as review of internal billing utilization and payment data to identify potential errors.

The goals of the data analysis program are to identify provider billing practices and services that pose the greatest financial risk to the Medicare program.

The MACs and UPICs shall document the processes used to implement their data analysis program and provide the documentation upon request.

In order to implement a data analysis program, the MACs and UPICs shall:

- Collect data from sources such as:

  o Historical data, e.g., review experience, denial data, provider billing problems, provider cost report data, provider statistical and reimbursement (PS&R) data, billing data, payment data, utilization data, data from other Federal sources (e.g., QIO, other MACs, Medicaid); and

  o Common Working File (CWF)

- Referrals from internal or external sources (e.g., 1-800 Medicare Call Center, provider audit, beneficiary, State Senior Medicare Patrol, or other complaints).

The shared system maintainer shall allow MACs the ability to select claims using the NPI or the legacy number (OSCAR or UPIN) as a criterion for medical review.

C. Resources Needed for Data Analysis

The MACs and UPICs shall have available sufficient hardware, software, and personnel with analytical skills to meet requirements for identifying problems efficiently, and effectively developing and implementing corrective actions. If MACs are unable to employ staff with the qualifications necessary for effective data analysis, evaluation and reporting, they shall use other entities (e.g., universities, consultants, other contractors) who can provide the technical expertise needed. The following are minimum resource requirements for conducting data analysis, evaluation, and reporting.

1. Data Processing Hardware

Adequate equipment for data analysis includes facilities to process data (e.g., mainframes and personal computers) and to store data (e.g., tape drive, disk drives, etc.). Upgrading current resources (e.g., mainframe computers, shared systems, etc.) or the purchase of new capabilities (e.g., microcomputer workstations or subcontracts for computer services) may provide additional processing capabilities. In addition, MACs and UPICs shall have secure telecommunication capabilities to interact with the CMS Data Center.

2. Data Processing Software

The CMS provides MACs and UPICs with software to allow communication

with the CMS Data Center. At their discretion, MACs and UPICs that wish to develop or acquire additional software that allows for analysis of internal data or other data obtained from the CMS Data Center may do so. The MACs and UPICs should have internal software to support the analyses of data to meet program goals.

3. **Personnel**

The MACs and UPICs shall have staff with appropriate training, expertise and skills to support the application of software and conduct systematic analyses and clinical evaluation of claims data. CMS strongly encourages MACs and UPICs to have staff with clinical expertise (e.g., registered nurses) and a mix of skills in programming, statistics, and data mining analysis (e.g., trending and profiling of providers/codes).

The MACs and UPICs shall also employ a staff that has training in developing analytical and sampling strategies for overpayment projections.

D. Frequency of Analysis

The MACs shall have a minimum of 18 months of data but are encouraged to have 36 months. The MACs shall, at a minimum, compare the current 6-month period to the previous 6-month period to detect changes in providers' current billing patterns and to identify trends in new services. Summary data or statistically representative samples can be used when dealing with very large volumes of data.

E. Determine Indicators to Identify Norms and Deviations

The MACs and UPICs shall develop indicators that will be used to identify norms, abnormalities, and individual variables that describe statistically significant time-series trends and the most significant abnormalities or trends. Examples of indicators or variables are:

- Standard deviations from the mean;

- Percent above the mean or median;

- Percent change in billing activity, payment charges, and number of visits/services from one period to another.

- Rate of change over specified periods in time.

F. Document Data Strategy

While the CMS is deliberately not prescriptive in terms of the technical details of how to reach data analysis goals, MACs and UPICs are expected to develop the most

sophisticated and effective methods and procedures to meet these goals and will be held accountable for accurate, effective reports, procedures, and quality outcomes.

2.3 – **Sources of Data for UPICs**
(Rev. 10365; Issued: 10-02-20; Effective: 08-27-20; Implementation: 08-27-20)

The term Medicare beneficiary identifier (Mbi) is a general term describing a beneficiary's Medicare identification number. For purposes of this manual, Medicare beneficiary identifier references both the Health Insurance Claim Number (HICN) and the Medicare Beneficiary Identifier (MBI) during the new Medicare card transition period and after for certain business areas that will continue to use the HICN as part of their processes.

A.   Contractors To Which This Section Applies

This section applies to UPICs.

B.   General

The UPICs approach for combining claims data (MAC data, Recovery Auditor data from the Recovery Auditor data warehouse) and other data to create a platform for conducting complex data analysis shall be documented in their Information Technology Systems Plan. By combining data from various sources, the UPIC will present an entire picture of a beneficiary's claim history regardless of where the claim was processed. The primary source of this data will be the CMS shared systems data, National Claims History (NCH), and Integrated Data Repository (IDR). The UPIC shall be responsible for obtaining data for all beneficiaries for whom the MAC(s) paid the claims.

At a minimum, UPICs are required to store the most recent 36 months' worth of data (including Part A, Part B, DME, home health & hospice) for the jurisdiction or zone defined in their task order.

If the jurisdiction of the MAC(s) is not defined geographically, the UPIC shall obtain a complete beneficiary claims history for each unique beneficiary for whom the MAC(s) paid a claim.

EXAMPLE 1: The MAC(s) jurisdiction covers Maryland but includes a hospital chain with facilities in Montana. The UPIC would request claims history from shared systems, NCH, or IDR for all claims paid by the MAC(s).

EXAMPLE 2: The MAC(s) jurisdiction covers Maryland, a beneficiary lives in Pennsylvania, and the beneficiary saw a doctor in Maryland. The UPIC would request from shared systems, NCH, or IDR for all claims paid by the MAC(s).

The UPICs will not be able to tap data from the Common Working File (CWF).

The UPICs should, at their discretion, if agreement and cooperation of the MAC(s) are obtained, use data directly from the claims processing system of the MAC(s), and then supplement the other data using NCH.

In developing this plan, the UPICs shall address the above requirements and, at a minimum, establish read-only access to the MAC's shared claims processing system(s) and access to the Part A, B, and D data available through the NCH for the jurisdictional area defined in the Task Order. The UPIC shall obtain denial data through the MACs and document the process for obtaining this data from the MAC(s) in the Joint Operating Agreement. At a minimum, the denial data shall include data for edits that were requested and/or recommended by the UPIC.

The UPIC shall have the ability to receive, load, and manipulate CMS data. The data shall also be maintained in accordance with CMS and Federal privacy laws and regulations as described in the CMS Data Use Agreement. For planning purposes, the UPICs should assume that there are 30 claims per Medicare beneficiary identifier (Mbi) per year, on average. A claim record is about 1000 bytes. To calculate the storage space necessary, use the following formula:

$$(\#Mbis) \times (30 \text{ claims}) \times (\#years) \times (1000) = \#bytes$$

The CMS contract officer's representative (COR) and UPIC will need to complete:

- A data use agreement to give permission to receive privacy protected data;

- A Data request form to specify all data required by the UPIC;

- A HDC application for HDC access and/or CMS systems' access to get access to the data center and/or to specify which CMS systems the UPIC will access;

    - A DESY system application form. (This is provided to the UPIC post- award).

2.4 – Sources of Data for MACs and UPICs
(Rev. 12772; Issued: 08-09-24; Effective: 09-20-24; Implementation: 09-20-24)

A. Contractors To Which This Section Applies

This section applies to MACs and UPICs. The sources of data for CERT and Recovery Auditors are specified in their SOWs.

B. General

The data sources that MACs and UPICs use will depend upon the issue(s) being addressed and the availability of existing data. CMS maintains numerous systems housing Medicare, Parts A, B, and D claims, Beneficiary Entitlement, Enrollment and

Utilization data, Provider reference information. The IDR is the enterprise resource designed to house and unify the data from disparate systems to enable cross-cutting reporting and analysis. The IDR has been created with an aim toward reducing data redundancy, providing flexibility to satisfy changing business needs and serve as the relational data warehouse for core CMS data. The IDR provides the system platform and database structures which enable one store of data to meet the various needs of our MAC and UPIC community. The repository is leveraged by multiple reporting, analytical and operational production applications.

Systematic data analysis requires MACs and UPICs to have in place the hardware and software capability to profile providers in aggregate, by provider type, by common specialties among providers, or individually. Some of the provider information that should be used includes:

- Types of providers;

- Volume of business;

- Volume (or percentage) of Medicare/Medicaid patients;

- Prevalent types of services;

- Location;

- Relationships to other organizations;

- Types of ownership;

- Previous investigations by the UPIC;

- Size and composition of staff;

- Administrative costs;

- Claims history; and

- Other information needed to explain or clarify the issue(s) in question.

Where possible, the selection of providers should show a representative grouping, in order to accurately reflect the extent of program losses.

C. Primary Source of Data

Claims data is the primary source of information used to identify and target fraudulent, wasteful or abusive activities. Sources of claims data are:

- IDR--MACs and UPICs should utilize the reports accessible in the system platform and database structures. Reports include the following:

  o Claims Summary Information report (CSI) which used to be called Health Care Information System (HCIS),

  o Part B Analytics System Report (PBASR), which show comparative utilization ratios by code, MAC, and specialty,

  o Short Term Alternatives for Therapy Services (STATS) which shows Outpatient therapy professional and provider claims data,

  o IDR Analysis Reports which include analysis reports and IDR volume and statistics, and

  o Focused Medical Review (FMR) which shows Part B claims utilization and enrollment data.

The MACs and UPICs shall also use national data where available. National data for services billed by skilled nursing facilities (SNFs) and home health agencies (HHAs) is available at the CMS Data Center. When made available, contractors can access through CMS One Program Integrity (see below One PI) or the CMS Enterprise Portal.

- CMS One Program Integrity (One PI) – Serves as a system application, tool and databases providing access to the CMS IDR;

- Contractor Local Claims Data – Local data should be compiled in a way to identify which providers or type of service in the contractor's area may be driving any unusual utilization patterns;

- CMS Fraud Prevention System(FPS)--When access is available, MACs should consider periodically reviewing the information and data trends resulting from national predictive models contained in the FPS;

- CMS PBASR--The Report stores data sets that contain annual timeframes, and Healthcare Common Procedure Coding System HCPCs/CPT codes that correspond to provider/supplier disciplines. Each data set displays the allowed services, allowed charges, and payment amounts by HCPCs/CPT codes and prominent modifiers. The PBAR is only accessible through the Enterprise Portal; and

- CMS Claims Summary Information (CSI)——Files contain Medicare Part A (i.e., Inpatient, Skilled Nursing Facility, Home Health Agency (Part A & B) and Hospice) and Medicare Part B (i.e., Outpatient) information based on the type and State of the institutional provider. The data set names correspond with

the provider type. Brief descriptions of the provider types and the selected reporting elements (e.g., units of service, billed charges, provider ZIP code, etc.) are provided. Access is through the One PI portal.

D. Secondary Sources of Data

The MACs and UPICs should consider other sources of data in determining areas for further analysis. These include:

- OIG and Government Accountability Office (GAO) reports;
- Fraud Alerts;
- Beneficiary, physician and provider complaints;
- Appeals data from QICs, including appeals overturn rate for a particular type of claim;
- Referrals from the QIO, other contractors, CMS components, Medicaid fraud control units, Office of the U.S. Attorney, or other federal programs;
- Suggestions provided directly or implicit in various reports and other materials produced in the course of evaluation and audit activities, (e.g., contractor evaluations, State assessment, CMS-directed studies, contractor or State audits of providers);
- Referrals from medical licensing boards;
- Referrals from the CAC;
- Peer Review Reports such as the First-look Analysis Tool for Hospital Outlier Monitoring (FATHOM) and Program to Evaluate Payment Patterns Electronic Report (PEPPER), and Comparative Billing Reports;
- Information on new technologies and new or clarified benefits;
- Provider cost reports;
- Provider Statistical and Reimbursement (PS&R) System data;
- Enrollment data;
- Overpayment data;
- Pricing, data analysis, and coding (PDAC) data;
- Referrals from other internal and/or external sources (e.g., MAC audit staff, audit staff or, MAC quality assurance (QA) staff);
- Medicare Learning Network – which includes MedLearn Matters articles and Quarterly Provider Compliance Newsletters;
- IBM Cognos support for the Part D and Drug Data Processing System (DDPS) using the Teradata data repository;
- CMS prepared data, such as a listing of distinct providers or suppliers and/or bills that require medical review.

While the MAC, Recovery Auditor, and UPIC should investigate reports from the GAO, congressional committees, Office of Inspector General Office of Audit Services (OIG OAS), OIG OI, newspaper and magazine articles, as well as local and national television and radio programs, highlighting areas of possible abuse, these types of leads should not *be used as a main* source for leads on fraud, waste or abuse cases.

Transmittals Issued for this Chapter

Rev #	Issue Date	Subject	Impl Date	CR#
R12772PI	08/09/2024	Updates of Chapter 1, Chapter 2, Chapter 3, Chapter 4, and Chapter 9 in Publication (Pub.) 100-08, Including Complaint Referral Coordination Between Contractors	09/20/2024	13719
R10365PI	10/02/2020	Updates to Chapters 1, 2, 3, 4, 6, 7, 8, 10, 11, and Exhibits of Publication (Pub.) 100-08	08/27/2020	11884
R10228PI	07/07/27/2020	Updates to Chapters 1, 2, 3, 4, 6, 7, 8, 10, 11, and Exhibits of Publication (Pub.) 100-08 Rescinded and replaced by Transmittal 10365	08/27/2020	11884
R876PI	04/12/2019	Update to Publication (Pub.) 100-08 to Provide Language-Only Changes for the New Medicare Card Project	05/13/2019	11109
R658PI	06/22/2016	Medicare Program Integrity Data Analysis--Update	02/01/2016	9176
R630PI	12/31/2015	Medicare Program Integrity Data Analysis—Update – Rescinded and replaced by Transmittal 658	02/01/2016	9176
R313PI	11/20/2009	Program Integrity Manual (PIM) Reorganization Chapters 1, 2, and 7	12/21/2009	6546
R279PI	12/19/2008	Zone Program Integrity Contractor (ZPIC) Updates	01/26/2009	6171
R231PI	01/04/2008	NPI Number for Medical Review	04/07/2008	5761
R180PI	12/22/2006	Sources of Data for PSCs	01/22/2007	5412
R101PI	01/28/2005	Benefit Integrity (BI) PIM Revisions	02/28/2005	3579
R071PI	04/09/2004	Rewrite of Program Integrity Manual (except Chapter 10) to Apply to PSCs	05/10/2004	3030
R047PI	07/25/2003	Data Analysis	08/08/2003	2517
R032PI	10/25/2002	Fraud Alerts	10/25/2002	2333
R016PIM	11/28/2001	Adds Various Program Memoranda for BI Requests for Information, Organizational Requirements, Unsolicited Voluntary Refund Checks, Anti-Kickback Statute Implications	11/28/2001	1732
R003PIM	11/22/2000	Complete Replacement of PIM Revision 1.	NA	1292

Rev #	Issue Date	Subject	Impl Date	CR#
R001PIM	06/2000	Initial Release of Manual	NA	931

Back to top of chapter