

# Medicare Provider Utilization and Payment Data

## Contents

### Database

#### List of Aging Related Databases

- [AHA \(America Hospital Association\)](#)

- [American Hospital Association Annual Survey of Hospitals](#)

- [American Medical Association Physician Master File](#)

- [Area Health Resource File\(AHRF\)](#)

- [Area Resource File \(ARF\)](#)

- [Behavioral Risk Factor Surveillance System\(BRFS\)](#)

- **Topic/focus:** Summarizes the utilization and payments for procedures, services, and prescription drugs provided to Medicare beneficiaries by specific inpatient and outpatient hospitals, physicians, and other suppliers
- **Year(s):** 2012, 2013, (some from 2011)
- **Source:**<http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/index.html>
- **Study and sample characteristics:** These data include information for the 100 most common inpatient services, 30 common outpatient services, all physician and other supplier procedures and services, and all Part D prescriptions. Providers determine what they will charge for items, services, and procedures provided to patients and these charges are the amount the providers bill for an item, service, or procedure.
- **Universe:**
- **Variables:**
  - Key Data Sources

(Extracted from the Methodology documents)  
The primary data source for these data is CMS's National Claims History (NCH) Standard Analytic Files (SAFs) which include claims with dates of service within the reported calendar year and accreted to the NCH as of the June following the reported calendar year. The NCH SAFs contain 100 percent of Medicare final action claims for beneficiaries who are enrolled in the FFS program. The NCH contains a SAF for each type of Medicare claim type including institutional (i.e., hospital inpatient, hospital outpatient, skilled nursing, home health and hospice) and non-institutional (i.e., Part B physician/supplier and DMEPOS). Specifically, the Part B Physician/Supplier SAF was used to create the Physician and Other Supplier PUF, which includes services from physicians, non-physician practitioners, laboratories, imaging, ambulances, etc. (does not include claims from the DMEPOS SAF). Beneficiary and service counts, provider charges, Medicare allowed amounts and payments, place of service, provider type, and Medicare participation indicator were summarized from this SAF.

- Population

(Extracted from the POS Methodology document) The Physician and Other Supplier PUF includes data for providers that had a valid

Chronic condition data warehouse

•

Dartmouth Atlas of Health Care

•

Dartmouth Atlas Project

•

Demographic Health Survey (DHS)

>

Healthcare Co and Utilization Project(H-CUP/HCUPNE

•

Health United States

•

ICPSR

•

Innovation in Medical Evidence Development & Surveillance(I

•

Integrated Public Use Microdata Series

•

Kaiser family foundation state level health facts

•

LSOA II

•

Medical Expenditure Panel Survey(MEPS)

•

NPI and submitted Medicare Part B non-institutional claims (excluding DMEPOS) during the 2012 and 2013 calendar years. To protect the privacy of Medicare beneficiaries, any aggregated records which are derived from 10 or fewer beneficiaries are excluded from the Physician and Other Supplier PUF.

NOTE: Records where the aggregation of claims less than **10** beneficiaries are NOT in the data set (Extracted from the Part D Methodology document) Information in the Part D Prescriber PUF is based on data from 35.7 million beneficiaries enrolled in the Medicare Part D prescription drug program, who comprise 68 percent of all Medicare beneficiaries. Approximately 23 million of these Part D beneficiaries are enrolled in stand-alone Prescription Drug Plans (PDP) and 13 million enrolled in Medicare Advantage Prescription Drug (MAPD) plans. NOTE: Records where the aggregation of claims less than **11** beneficiaries are NOT in the data set (for either all or ge65 subset)

- Data Limitations

These data sets are not representative of a provider's entire practice or operation; the data only includes information relating to Medicare beneficiaries with Part B FFS coverage. This data also does not include any other information for procedures not submitted to medicare (such as individuals with other insurance coverage, or uninsured individuals). HCPCS modifier codes are not included in this data set, so that the aggregated values for each provider's billed HCPCS code will include the total for all modifier variations.

- Methodology Files:

- [Medicare Physician and Other Supplier PUF Methodology.pdf \(S3\)](#)
- [Medicare Provider Charge Data CY2013 Inpatient Methodology.pdf \(S3\)](#)
- [Medicare Provider Charge Data CY2013 Outpatient Methodology.pdf \(S3\)](#)
- [Medicare Provider Charge Data CY2013 Part D Prescriber Methods.pdf \(S3\)](#)

- Limitations

- The data is aggregated by each provider (about 1.3 million providers) and by the charge code unique to that data set. So in the case of the POS data set, each provider will have one entry for each unique HCPCS code that provider billed for in the data set time window (calendar year 2013). The same aggregation occurs for Part D prescriber where each prescriber has an entry for each unique medication that was prescribed during the calendar

Medicare Current Beneficiary Survey (MCBS)

•

Medicare Healthcare Cost Report Information System(HCRIS)

•

**Medicare Provider Utilization and Payment Data**

•

MEPSnet

•

National Ambulatory Medical Care Survey (NAMCS)

•

National Health Aging Trends Study(NHATS)

•

National Health and Nutrition Examination Survey (NHANES)

•

National Health Examination Survey(NHES) Cycle\_II

•

National Health Examination Survey (NHES) Cycle I

•

year 2013. For Outpatient data, each provider has one entry for each APC code that was billed.

- Unfortunately, the inpatient table has a DIFFERENT time interval. Instead of the calendar year 2013, inpatient data is aggregated over the **FISCAL** year 2013. So you can't really relate inpatient and outpatient this way.
- Data Import Normalization
  - Provider Normalization
    - All NPPES data (from POS and PartD) are merged into a common NPPES data table. All unique fields are maintained. The only addition is that for providers that were merged from the POS data set, the provider\_type\_flag is set to 'O' (which is a new value for that field assuming that the original field only supported 'S' and 'T').
    - Provider Type string was normalized into a separate table to reduce the data size as there are only 246 provider types, but those strings are repeated for all 1.3 million providers.
    - The State code was patched for Army Addresses in foreign countries where the zip code matched the standard USPS 9 digit format. The state codes are patched to AE, AP, and AA based on the original country code.
    - Zip codes are all normalized to 5 digit or 9 digit zip codes without spaces or dashes.
  - Inpatient and Outpatient Provider Normalization
    - The two provider data sets are merged into one data set because there is an overlap and the fields provided by each data set are identical.
  - APC and DRG Codes
    - The original data has the code and description in a single string. These values have been split apart so that the short code (382 or A834) is used to join the code table to the main inpatient or outpatient table. This also allows for quickly filtering by APC or DRG code.
  - HRR Code
    - The HRR code was originally formatted as "SS - Region". The state and region strings are separated into separate fields for easier filtering.
  - Drug names

National Health Examination Survey (NHES) Cycle III

•

National Health Interview Survey(NHIS)

•

National Health Interview Survey (NHIS) Multiple Cause-of-Death Public Use Data File

•

National Health Interview Survey (NHIS) State Data Files

•

National Health Interview Survey on Disability(NHI: D) Phase I and II

•

National Home and Hospice Care Survey (NHHCS)

•

National Hosp Ambulatory Medical Care Survey(NHAM)

•

National Hospital Discharge Survey(NHDS)

•

- The brand drug names and generic drug names were normalized out into a separate table for better storage and easier searching. When filtering by name, remember to join these tables back to the main partd data set.
- Numeric Values
  - Where identified, the data sets were fully scanned and an appropriate numeric storage type was identified for each field. All numeric values are either INTEGER or DECIMAL(x,y). When using derived or calculated fields, remember to cast the number type to the type you are expecting using PostgreSQL formatting (such as ::DECIMAL(13,2))
- **Access:** Publicly available
- **Cost:** Free
- **Key web links:**
  - [Main Page](#)
  - [Physician and Other Suppliers](#)
    - Database normalized and loaded into Amazon Redshift in the sandbox/medicare2013 database
    - Tables Provided by this data set:
      - POS - The Provider and Other Suppliers billing activity by provider and by unique HCPCS code
      - HCPCS - The HCPCS codes and their descriptions
      - PLACE - A Place of Service table as imported from the methodology document
      - NPPES - (Shared by Part D) National Plan and Provider Enumeration System derived Provider information for all providers appearing in the POS and PartD data sets.
      - NPPES\_TYPE - (Shared by Part D) The Provider specialty type normalized out (generally informational only)
  - [Inpatient](#)
    - Tables Provided by this data set:
      - IOPROVIDER - (Shared with Outpatient) The Provider Information table (facility, location, etc.)
      - HRR - (Shared with Outpatient) The provider's assigned Hospital Referral Region (includes state and region name breakdown)
      - DRG - Medicare severity Diagnostic Related Group code for inpatient

National Nursing Home Survey(NNHS)

•

National Survey of Ambulatory Surgery (NSAS)

•

Nationwide Inpatient Sample(NIS)

•

NCHS-CMS Medicare and NCHS-USRDS Data

•

NDFRT

>

RESDAC CMS Data

•

School Health Policies and Programs Study(SHPPS)

•

SEER-Medicare

•

SIPP

•

State of Aging and Health Report

•

The biennial National Study of Long-Term Care Providers (NSLTCP)

•

The MIMIC II Clinical Database Demo

•

procedures broken down into code and description

- INPATIENT - Inpatient charges by provider and DRG code including submitted, charged, and paid values

- **Outpatient**

- Tables Provided by this data set:
  - IOPROVIDER - (Shared with Outpatient) The Provider Information table (facility, location, etc.)
  - HRR - (Shared with Outpatient) The provider's assigned Hospital Referral Region (includes state and region name breakdown)
  - APC - Ambulatory Procedure Classification code for outpatient procedures broken down into code and description
  - OUTPATIENT - Outpatient charges by provider and APC code including submitted, charged, and paid values

- **Part D Prescriber**

- Tables Provided by this data set:
  - PARTD\_PDSUM - Part D Providers billing summary including aggregation data not provided in the Part D data set
  - PARTD - Part D utilization by prescriber and drugname including costs, claims, days supplied, etc. Some data separation provided for claims for beneficiaries 65 or older
  - DRUGNAMES - Brand drug names normalized out and indexed by a unique drug number (unique to this database)
  - GENERICNAMES - Generic drug names normalized out and index by a unique drug number (unique to this database)
  - NPPES - (Shared by POS) National Plan and Provider Enumeration System derived Provider information for all providers appearing in the POS and PartD data sets.
  - NPPES\_TYPE - (Shared by POS) The Provider specialty type normalized out (generally informational only)

- **Summary:**

## Sources

The National  
Hospital Care  
Survey(NHCS)

•

The  
Organisation  
for  
Economic  
Co-operation  
and  
Development  
international  
health  
statistics

•

U.S.  
Census  
Bureau  
Data

•

U.S.  
Nursing  
Home  
Information  
and  
Registry

•

Us  
Preventive  
Services  
Task Force

•

Visiting  
Nurse  
Service of  
New York

•

Web-Based Inj  
Statistics Que  
and Reporting  
System(WISQ,

•

Wide-Ranging  
Data for Epide  
Research(CDC

•

ZCTA  
County

•

Download,  
normalization,  
and upload  
procedure for  
datasets.

All information relating to data elements descriptions and database summaries have been pulled from the database providers website located here: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/index.html>. Summary information has been pulled from <http://www.mcw.edu/edrc/datacatalog.htm>.

•

In-Progress  
Databases

•

LastPass  
Enterprise

>

Software Tools

>

Training  
Materials