

SDRC Tip Sheet – Public Use Files

The State Data Resource Center (SDRC) team compiled this document highlighting free additional datasets that state Medicaid agencies can use to better understand the Medicare enrollee experience in their state and nationally.

Table 1. Public Use Files (PUFs)

PUF	Description
Medicare-Medicaid Linked Enrollee Analytic Data Source (MMLEADS) Version 2.0 MMLEADS PUF User Guide	<p>The MMLEADS PUF data are based on MMLEADS version 2.0. The 2006–2012 aggregated data include information on full and partial Medicare–Medicaid enrollees as well as, for comparison purposes, Medicare-only and Medicaid-only (eligible via disability) enrollees. The MMLEADS PUF data and user guide are available for download.</p> <p>CONTAINS: Dual status data.</p>
RxNorm	<p>RxNorm provides normalized names for clinical drugs and links their names to many of the drug vocabularies commonly used in pharmacy management and drug interaction software, including those of First Databank, Micromedex, Medi-Span, Gold Standard Drug Database, and Multum. By providing links between these vocabularies, RxNorm can mediate messages between systems that do not use the same software and vocabulary. RxNorm now includes the National Drug File-Reference Terminology (NDF-RT) from the Veterans Health Administration. NDF-RT is used to code clinical drug properties, including the mechanism of action, physiologic effect, and therapeutic category.</p>
U.S. Department of Veterans Affairs (VA) National Drug File	<p>The VA National Drug File includes information on generic prescription drugs. The information pertains to manufacturers, drug classification, dosage, intake methodology, etc.</p>
Medicare Provider Utilization and Payment Data	<p>The Centers for Medicare & Medicaid Services (CMS) has released a series of publicly available data files that summarize the utilization of and payments for procedures, services, and prescription drugs provided to Medicare beneficiaries by specific inpatient and outpatient hospitals, physicians, and other suppliers. These Medicare Provider Utilization and Payment Data files include information for common inpatient and outpatient services, all physician and other supplier procedures and services, and all Part D prescriptions. The SDRC team believes these files may be good alternatives to working with the claims data.</p>

PUF	Description
Healthcare Cost Report Information System (HCRIS)	<p>HCRIS contains annual reports submitted by institutional providers to Medicare. It provides information to CMS that assists with the annual settlement summary between CMS and the institutional provider. The cost report information includes facility-level utilization statistics, costs, and charges by cost center (in total and for Medicare).</p>
Provider of Services (POS) File	<p>For POS files since the second quarter of 2011 and all future POS files, the POS Extract is created from the Quality Improvement Evaluation System database. The file contains an individual record for each Medicare-approved provider and is updated quarterly. The file includes information for all institutional providers, ambulatory surgical centers, and clinical laboratories.</p> <p>CONTAINS: Provider-level data</p>

PUF	Description
<p>National Provider Identifier (NPI) Downloadable File</p>	<p>The NPI is a unique identification number for covered healthcare providers. The NPI is a 10-position, intelligence-free numeric identifier (10-digit number); this means that the numbers do not carry other information about healthcare providers, such as the state in which they live or their medical specialty. The key data elements that are found in the download include the following:</p> <ul style="list-style-type: none"> • NPI. • Entity Type Code (1-Individual or 2-Organization). • Provider Name (first name, middle name, last name, prefix, suffix, credential(s); OR the legal business name for organizations). • Provider Other Name (first name, middle name, last name; OR “doing business as” name, former legal business name, other name for organizations). • Provider Business Mailing Address (first-line address, second-line address, city, state, postal code, country code if outside the United States, telephone number, fax number). • Provider Business Location Address (first-line address, second-line address, city, state, postal code, country code if outside the United States, telephone number, fax number). • Health Care Provider Taxonomy Code(s). • Other Provider Identifier(s). • Other Provider Identifier Type Code. • Provider Enumeration Date. • Last Update Date. • NPI Deactivation Date. • NPI Reactivation Date. • Provider Gender Code. • Provider License Number. • Provider License Number State Code. <p>For more information or to download this file, please visit the CMS website.</p> <p>CONTAINS: State-level data</p>
<p>Medicare Part B Procedure Counts by State</p>	<p>This ZIP file includes Part B data categorized by procedure codes for each state/territory.</p> <p>CONTAINS: State-level data</p>

Table 2. Accountable Care Organization (ACO)-Related Analyses PUFs

ACO-Related Analyses PUF	Description
Shared Savings Program ACOs Provider-Level Research Identifiable File (RIF)	<p>The CMS Center for Medicare (CM) has created a set of standard analytical files that contain Shared Savings Program ACO data.</p> <ul style="list-style-type: none"> • Contains ACO-specific metrics as well as summarized beneficiary and provider information for each performance year of the Shared Savings Program. • Data available for 2013 to 2016. <p>CONTAINS: Provider-level data</p>
Shared Savings Program ACOs Beneficiary-Level RIF	<p>CM has created a set of standard analytical files that contain Shared Savings Program ACOs.</p> <ul style="list-style-type: none"> • The beneficiary-level RIF contains a record for every beneficiary who participated in the Shared Savings Program. <p>The data file is structured in a way that allows CMS to respond to a wide range of potential inquiries. For example, analysts should be able to use the data to examine:</p> <ul style="list-style-type: none"> • Differences in beneficiary characteristics between ACO and non-ACO beneficiaries. • Stability of beneficiary alignment during and across performance years and the extent to which aligned beneficiaries receive care outside of the ACO. • Differences in beneficiary characteristics between ACO-aligned beneficiaries who were and were not used for financial reconciliation. <p>The Shared Savings Program ACO file does not include Pioneer, Next Generation, or Comprehensive End-Stage Renal Disease (ESRD) Care ACOs.</p>
Pioneer ACOs RIF	<p>CMS has created a set of standard analytical files that contain Pioneer ACOs:</p> <ul style="list-style-type: none"> • Provider-level RIF that contains a record for each individual provider that participated in the Pioneer ACO Model. • Beneficiary-level RIF that contains a record for each individual beneficiary that was aligned to a Pioneer ACO. • Pioneer ACO Settlement File. <p>CONTAINS: Beneficiary- and provider-level data</p>

Table 3. Provider Use Files

Provider Use File	Description
National Plan and Provider Enumeration System (NPPES) Registry	<p>The downloadable file is available here. After CMS supplies an NPI, CMS publishes the parts of the NPI record that have public relevance, including the provider’s name, specialty (taxonomy), and practice address.</p> <p>CONTAINS: Provider-level data</p>
Hospital Compare Tool	<p>A geographic search function allows users to view individual hospitals, related details, and that facility’s performance on various quality measures.</p> <p>CONTAINS: Provider-level data</p>
Nursing Home Compare Tool	<p>A geographic search function allows users to view individual nursing homes, related details, and that facility’s performance on various quality measures.</p> <p>CONTAINS: Provider-level data</p>
Home Health Compare Tool	<p>A geographic search function allows users to view individual home health agencies, related details, and that entity’s performance on various quality measures.</p> <p>CONTAINS: Provider-level data</p>
Physician Compare Tool	<p>A geographic search function allows users to view individual health care providers/group practices, demographic details, and group practice-level performance on various quality measures.</p> <p>CONTAINS: Provider-level data</p>

Provider Use File	Description
<p>Physician Compare Data Dictionary</p>	<p>CMS provides official datasets for the Medicare.gov Physician Compare website to give useful information about group practices, physicians, and other healthcare professionals currently enrolled in Medicare.</p> <p>The Physician Compare National Downloadable File includes general information about individual eligible clinicians (ECs), such as demographic information and Medicare quality program participation.</p> <p>The Physician Compare 2017 Individual EC Public Reporting – Clinical Quality of Care file contains performance rates for the six Individual EC 2017 Physician Quality Reporting System (PQRS) measures reported via claims that are available for public reporting. EC measure data are available for 86,053 individual ECs.</p> <p>The Physician Compare 2017 Group Practice Public Reporting – Clinical Quality of Care file contains performance rates for the 125 group practice 2017 PQRS measures reported via the web interface that are available for public reporting. Group practice measure data are available for 13,083 group practices.</p> <p>The Physician Compare 2017 Group Practice Public Reporting – Patient Experience file contains performance rates for the eight Consumer Assessment of Healthcare Providers and Systems (CAHPS) for PQRS measures reported by group practices that are available for public reporting. CAHPS for PQRS measure data are available for 239 group practices.</p> <p>Because only a subset of the measures is included, not all healthcare professionals and group practices that satisfactorily participated in PQRS and submitted quality measures will have measure data in these files. For more information about which measures are available for public reporting, visit the Physician Compare Initiative page.</p> <p>Note: Because of data use agreements with data vendors, not all data on Physician Compare can be shared in the Physician Compare National Downloadable File. For more information about what is included in this database and how it differs from the information on the Physician Compare website, refer to the Downloadable Database Dictionary (under Get Supporting Documents).</p> <p>CONTAINS: Provider-level data</p>

Provider Use File	Description
Dialysis Facility Compare Tool	<p>Medicare has data you can use to compare dialysis facilities (centers) based on the quality of patient care they provide. You can also compare their patient experience survey results.</p> <p>CONTAINS: Provider-level data</p>
Supplier Directory	<p>Allows users to search for suppliers by ZIP Code to find suppliers of durable medical equipment, prostheses and prosthetic devices, orthotics, and supplies.</p> <p>CONTAINS: Provider-level data</p>
CMS Compare Sites' Downloadable Datasets	<p>This website provides links on various comparison downloadable dataset websites for hospitals, nursing homes, physicians, home health agencies, dialysis facilities, hospices, inpatient rehabilitation facilities, long-term care hospitals, and suppliers.</p> <p>CONTAINS: Provider-level data</p>
Medicare Cost Reports	<p>Facilities must file CMS cost reports at the end of each fiscal year detailing items such as facility size (number of beds), profit status, and amount of care to indigent patients.</p>
Medicare Provider Utilization and Payment Datasets	<p>Provides summary data on which physicians receive payment for services from Medicare and what services they provided.</p> <p>CONTAINS: Provider-level data</p>
Open Payments	<p>Documents payments to physicians by pharmaceutical companies and other organizations. You can access both the user interface and the downloadable version.</p>

Table 4. Basic Stand Alone (BSA) Files

BSA File	Description
CMS Data Entrepreneurs' Synthetic Public Use File (DE-SynPUF)	<p>This is a synthetic data file with a structure similar to the Limited Data Sets files, but imputed values are used instead of real claims data. The file is intended to be used for training and data exploration by those who do not yet have access to real claims data. It is derived from the claims data, but the data have been altered so much that the file is not useful for research purposes. However, it could be useful for state Medicaid agencies that are considering obtaining Medicare claims data by allowing them to get a feel for what it is like to work with the data. The DE-SynPUF is available for 2008 to 2010 and includes beneficiary summary, inpatient, outpatient, carrier, and Part D Prescription Drug Event (PDE) data.</p>
Chronic Conditions BSA PUF	<p>The PUF is stratified by dual eligibility but is not at the beneficiary level. It contains one record for each combination of sex, age category, various chronic conditions indicators, and dual eligibility status. Counts of beneficiaries in the category as well as average payment amounts are available. The other BSA files do not have any dual eligibility information, and it is not possible to merge them with each other or with other data sources.</p> <p>CONTAINS: Dual-eligibility status data.</p>
CMS Medicare and Medicaid Statistical Supplement	<p>The supplement does not include personally identifiable information (PII) or individual-level data but rather total counts of Medicare individuals in Medicare fee-for-service (FFS) and Medicare Advantage (MA). This data was published annually in electronic form from 2001 to 2013.</p>
CMS Cost Report	<p>The report does not include PII or individual-level data but rather contains provider information, such as facility characteristics, utilization data, costs and charges (total and for Medicare), and other financial data. Data are reported at the facility level for hospitals, skilled nursing facilities, home health agencies, renal dialysis facilities, hospices, health clinics, and community mental health centers. Though data are reported every fiscal year (the most up-to-date information is as of 2019), they are not complete for all providers.</p> <p>CONTAINS: Provider-level data</p>

BSA File	Description
Kaiser Family Foundation Total Number of Medicare Beneficiaries Report	<p>The report does not include PII or individual-level data but rather the total count reported at the state and county levels, which comprises both Medicare FFS and MA counts from the penetration file. Further, the report has a lag time of approximately 2 years (e.g., as of 2020, the most recent report is from 2018).</p> <p>CONTAINS: State- and county-level data</p>
Claims BSA PUFs	<p>The BSA PUFs provide de-identified 5-percent samples of the 2008 claims used as the source for the Medicare utilization summary; exact dates and geographic information are not included.</p>
CMS Medicare Data on Geographic Variation	<p>Developed by CMS, these reports aim to provide information on geographic variation in the utilization and quality of healthcare services for the Medicare FFS and MA populations. Report data include summaries for all 50 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands at the state/territory and county levels (American Samoa, Guam, and the Northern Mariana Islands are not included). In addition, counts under 11 are suppressed. Despite data that includes 100 percent of individuals enrolled in Medicare FFS, there is a substantial data lag (e.g., as of 2020, the most recent update included 2008–2018 data). MA enrollment and participation data were added to the February 2015 update.</p> <p>CONTAINS: State- and county-level data</p>

Table 5. Visualization Tools

Visualization Tool	Description
Market Saturation and Utilization Data Tool	<p>The Market Saturation and Utilization Data Tool includes an interactive map and a dataset that shows national-, state-, and county-level provider services and utilization data for selected health service areas. The tool can be used to monitor market saturation as a means to prevent fraud, waste, and abuse. The data can also be used to reveal the degree to which use of a service is related to the number of providers servicing a geographic region.</p> <p>CONTAINS: National-, state-, and county-level data.</p>

Visualization Tool	Description
Centers for Disease Control and Prevention (CDC) Stats of the States	<p>CDC’s maps include state-by-state information on a variety of health indicators, such as cancer mortality and birth rates, and compare rates between states. The data comes from CDC’s birth and death reports.</p> <p>CONTAINS: State-level data.</p>