

# 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 for better understanding the Medicare enrollee experience in their state and nationally.

## Public Use Files (PUFs)

**Table 1. Public Use Files**

Public Use File	Description
<a href="#">Medicare-Medicaid Linked Enrollee Analytic Data Source (MMLEADS) Version 2.0</a> <a href="#">MMLEADS PUF User Guide</a>	<p>The MMLEADS PUF data are based off MMLEADS, Version 2.0. The 2006–2012 aggregated data include information on full and partial Medicare-Medicaid enrollees and—for comparison purposes—Medicare-only enrollees and Medicaid-only (eligible via disability) enrollees. The MMLEADS PUF data and user guide are available for download.</p> <p><b>CONTAINS: Dual-status data.</b></p>
<a href="#">RxNorm</a>	<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 not using the same software and vocabulary. RxNorm now includes the National Drug File – Reference Terminology (NDF-RT) from the Veterans Health Administration. NDF-RT is a terminology used to code clinical drug properties, including the mechanism of action, physiologic effect, and therapeutic category.</p>
<a href="#">U.S. Department of Veterans Affairs (VA) National Drug File</a>	<p>The VA’s National Drug File, which was updated in January 2018, includes information on generic prescription drugs. The information pertains to manufacturers, drug classification, dosage, intake methodology, etc.</p>

Public Use File	Description
<a href="#">Medicare Provider Utilization and Payment Data</a>	<p>The Centers for Medicare &amp; Medicaid Services (CMS) has released a series of publicly available data files that summarize the utilizations 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>
<a href="#">Healthcare Cost Report Information System (HCRIS)</a>	<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, chargers, Medicare payments, and financial information. For more information or to download these data, please visit the <a href="#">CMS website</a>.</p>
<a href="#">Provider of Services (POS) File</a>	<p>The POS Extract is created from the Quality Improvement Evaluation System database as of the second quarter of 2011 and all future POS files. 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><b>CONTAINS: Provider-level data.</b></p>

Public Use File	Description
<a href="#">NPI Downloadable File</a>	<p>The National Provider Identifier (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:</p> <ul style="list-style-type: none"> <li>• NPI.</li> <li>• Entity Type Code (1-Individual or 2-Organization).</li> <li>• Provider Name (first name, middle name, last name, prefix, suffix, credential(s), OR the legal business name for the organizations).</li> <li>• Provider Other Name (first name, middle name, last name, OR “doing business as” name, former legal business name, other name for organizations).</li> <li>• Provider Business Mailing Address (first-line address, second-line address, city, state, postal code, and country code if outside the United States, telephone number, fax number).</li> <li>• Provider Business Location Address (first-line address, second-line address, city, state, postal code and country code if outside the United States, telephone number, fax number).</li> <li>• Healthcare Provider Taxonomy Code(s).</li> <li>• Other Provider Identifier(s).</li> <li>• Other Provider Identifier Type Code.</li> <li>• Provider Enumeration Date.</li> <li>• Last Update Date.</li> <li>• NPI Deactivation Date.</li> <li>• NPI Reactivation Date.</li> <li>• Provider Gender Code.</li> <li>• Provider License Number.</li> <li>• Provider License Number State Code.</li> </ul> <p>For more information or to download this file, please visit the <a href="#">CMS website</a>.</p> <p><b>CONTAINS: State-level data.</b></p>
<a href="#">Medicare Part B Procedure Counts by State</a>	<p>This zip file includes Medicare Part B data categorized by procedure codes for each state/territory.</p> <p><b>CONTAINS: State-level data.</b></p>

Accountable Care Organizations (ACO)-Related Analyses

**Table 2. ACOs-Related Analyses PUFs**

Public Use File	Description
<p><a href="#">Shared Savings Program ACOs Provider-Level RIF</a></p>	<p>The CMS Center for Medicare (CM) has created a set of standard analytical files that contain Shared Savings Program ACOs.</p> <ul style="list-style-type: none"> <li>• Contains ACO-specific metrics as well as summarized beneficiary and provider information for each performance year of the Shared Savings Program.</li> <li>• Data available for years 2013–2016.</li> </ul> <p><b>CONTAINS: Provider-level data.</b></p>
<p><a href="#">Shared Savings Program ACOs Beneficiary-Level RIF</a></p>	<p>CM has created a set of standard analytical files that contain Shared Savings Program ACOs.</p> <ul style="list-style-type: none"> <li>• Beneficiary-level research identifiable file (RIF) that will contain a record for every beneficiary who participated in the Shared Savings Program.</li> </ul> <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"> <li>• Differences in beneficiary characteristics between ACO and non-ACO beneficiaries.</li> <li>• Stability of beneficiary alignment during and across performance years and the extent to which aligned beneficiaries receive care outside of the ACO.</li> <li>• Differences in beneficiary characteristics between ACO-aligned beneficiaries who were and were not used for financial reconciliation.</li> </ul> <p>The Shared Savings Program ACO file does not include Pioneer, Next Generation, or Comprehensive ESRD Care ACOs.</p>

Public Use File	Description
<a href="#">Pioneer ACOs RIF</a>	<p>CMS has created a set of standard analytical files that contain Pioneer ACOs.</p> <ul style="list-style-type: none"> <li>• Provider-level RIF that contains a record for each individual provider that participated in the Pioneer ACO Model.</li> <li>• Beneficiary-level RIF that contains a record for each individual beneficiary who was aligned to a Pioneer ACO.</li> <li>• Pioneer ACO Settlement File.</li> </ul> <p><b>CONTAINS: Beneficiary- and provider-level data.</b></p>

Provider User Files

**Table 3. Provider Use Files**

Provider Use File	Description
<a href="#">National Plan and Provider Enumeration System Registry</a>	<p>The downloadable file is available <a href="#">here</a>. 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><b>CONTAINS: Provider-level data.</b></p>
<a href="#">Hospital Compare Tool</a>	<p>Geographic search function that allows for viewing individual hospitals, related details, and that facility’s performance on various quality measures.</p> <p><b>CONTAINS: Provider-level data.</b></p>
<a href="#">Nursing Home Compare Tool</a>	<p>Geographic search function that allows for viewing individual nursing homes, related details, and that facility’s performance on various quality.</p> <p><b>CONTAINS: Provider-level data.</b></p>
<a href="#">Home Health Compare Tool</a>	<p>Geographic search function that allows for viewing individual home health agencies, related details, and that entity’s performance on various quality.</p> <p><b>CONTAINS: Provider-level data.</b></p>

Provider Use File	Description
<a href="#">Physician Compare Tool</a>	<p>Geographic search function that allows for viewing individual healthcare providers/group practices, demographic details, and group practice-level performance on various quality measures.</p> <p><b>CONTAINS: Provider-level data.</b></p>

Provider Use File	Description
<p><a href="#">Physician Compare Data Dictionary</a></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 professionals (EPs), such as demographic information and Medicare quality program participation.</p> <p>The Physician Compare 2014 Individual EP Public Reporting – Clinical Quality of Care file contains performance rates for the six Individual EP 2014 Physician Quality Reporting System (PQRS) measures reported via claims that are available for public reporting. EP measure data are available for 37,464 individual EPs.</p> <p>The Physician Compare 2014 Group Practice Public Reporting – Clinical Quality of Care file contains performance rates for the 14 group practice 2014 PQRS measures reported via the web interface that are available for public reporting. Group practice measure data are available for 271 group practices.</p> <p>The Physician Compare 2014 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 281 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><b>CONTAINS: Provider-level data.</b></p>

Provider Use File	Description
<a href="#">Dialysis Facility Compare Tool</a>	<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><b>CONTAINS: Provider-level data.</b></p>
<a href="#">Supplier Directory</a>	<p>Search for suppliers by ZIP Code to find suppliers of durable medical equipment, prostheses and prosthetic devices, orthotics, and supplies.</p> <p><b>CONTAINS: Provider-level data.</b></p>
<a href="#">CMS Compare Sites' Downloadable Datasets</a>	<p>This website provides links on various comparison downloadable dataset websites for hospitals, nursing homes, physicians, home health, dialysis facilities, hospices, inpatient rehabilitation facilities, long-term care hospitals, and suppliers.</p> <p><b>CONTAINS: Provider-level data.</b></p>
<a href="#">Medicare Cost Reports</a>	<p>Facilities must file CMS cost reports at the end of each fiscal year detailing items including facility size (number of beds), profit status, and amount of care to indigent patients.</p>
<a href="#">Medicare Provider Utilization and Payment Datasets</a>	<p>Summary data on which physicians receive payment for services from Medicare and what services they provided.</p> <p><b>CONTAINS: Provider-level data.</b></p>
<a href="#">Open Payments</a>	<p>Documents payments to physicians by pharmaceutical companies and other organizations. You can access both the <a href="#">user interface</a> and the <a href="#">downloadable version</a>.</p>



**Table 4. Basic Stand-Alone (BSA) Files**

Provider Use File	Description
<a href="#">CMS DE-SynPUF</a>	<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, so they can get a feel for what it is like to work with the data. The DE-Syn PUF is available for 2008–2010 and includes a beneficiary summary file, inpatient, outpatient, carrier, and Part D Prescription Drug Events.</p>
<a href="#">Chronic Conditions BSA PUF</a>	<p>The PUF is stratified by dual eligibility, but is not beneficiary level. It has 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><b>CONTAINS: Dual-eligibility status data.</b></p>
<a href="#">CMS Medicare and Medicaid Statistical Supplement</a>	<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. This information has a lag time of approximately 2 years (e.g., as of 2017, the most recent supplement is dated 2013).</p>
<a href="#">CMS Cost Report</a>	<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 2014), they are not complete for all providers.</p> <p><b>CONTAINS: Provider-level data.</b></p>

Provider Use File	Description
<a href="#">Kaiser Family Foundation Total Number of Medicare Beneficiaries Report</a>	<p>The report does not include PII or individual-level data, but rather the total count reported at the state and county levels, which is comprised of both Medicare FFS and Medicare Advantage counts from the penetration file. Further, the report has a lag time of approximately 2 years (e.g., as of 2017, the most recent report is from 2015).</p> <p><b>CONTAINS: State- and county-level data.</b></p>
<a href="#">Claims BSA PUFs</a>	<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>
<a href="#">CMS Medicare Data on Geographic Variation</a>	<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 Medicare Advantage 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). Further, counts under 11 are suppressed. Despite data including 100 percent of individuals enrolled in Medicare FFS, there is a substantial data lag (e.g., as of 2017, the most recent update was in February 2015, which included 2008–2013 data). Medicare Advantage enrollment and participation data were added to the most recent update of this PUF.</p> <p><b>CONTAINS: State- and county-level data.</b></p>

**Table 5. Visualization Tools**

Public Use File	Description
<a href="#">Market Saturation and Utilization Data Tool</a>	<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><b>CONTAINS: National-, state-, and county-level data.</b></p>

Public Use File	Description
<a href="#">CDC's Stats of the States</a>	<p>The Centers for Disease Control and Prevention's (CDC) maps include state-by-state information on a variety of health indicators, such as cancer mortality and birth rates, and then compare rates between states. The data comes from the CDC's birth and death reports.</p> <p><b>CONTAINS: State-level data.</b></p>