MAX
CENTERS FOR MEDICARE & MEDICAID SERVICES

MEDICAID POLICY BRIEF

MATHEMATICA Policy Research

Brief 21 • August 2015

Guide to MAX Data

Laura Ruttner, Rosemary Borck, Jessica Nysenbaum, and Susan Williams

The Medicaid Analytic eXtract (MAX) data system is the only national, person-level data set available for the Medicaid program. It includes extensive individual-level data on Medicaid enrollees and their Medicaid-financed health services. The research-ready data files are in a calendar-year format and can be used to answer a broad range of questions of interest to researchers, Medicaid officials, policymakers, and other stakeholders. A wealth of research has been conducted with the files.[1]

MAX is produced by the Centers for Medicare & Medicaid Services (CMS) and is derived from data submitted by states through the Medicaid Statistical Information System (MSIS). Annual MAX data files include enrollment and claims data for all Medicaid enrollees in the 50 states and the District of Columbia as well as for Medicaid-expansion CHIP (M-CHIP) enrollees.[2]

About This Issue Brief

The CMS Medicaid Analytic eXtract (MAX) data system contains a wealth of national and state-specific information about Medicaid enrollment, service use, and expenditures from 1999 through 2012. This issue brief is a resource for researchers interested in using MAX data. The brief provides an overview of the MAX data system, describes publicly available MAX data resources, and includes tips on using MAX data for research purposes.

MAX overview

CMS developed the MAX data system as a research tool for examining annual Medicaid enrollment, service use, and expenditures. MAX data files are available starting for 1999, the first year of MSIS data collection for all states. In MSIS, data are reported by federal fiscal year quarter for claims adjudicated for payment in that quarter. Given that provider billing may lag in time, MSIS data for a quarter do not include all services provided in that quarter. MAX takes the quarterly MSIS data files and transforms them into a more user-friendly calendar-year format that is based on date of service. Once released, MAX data are final. Researchers may use MAX to assess trends in enrollment, service use, and expenditures or patterns of care over time by linking individuals using permanent identification numbers over multiple years of MAX data. The MAX data system includes several enhancements from the MSIS data submitted by states:

- It combines initial claims, voided claims, and adjustments into final action claims.
- It maps information on some services from claims submitted in MSIS to new variables that provide additional detail that may be helpful to researchers.
- It accounts for retroactive enrollment and enrollment correction records that are reported in later quarters in MSIS.
- It separates Section 1915(c) waivers, which are reported as a single category in MSIS, into more detailed waiver-type categories.
- It corrects coding inconsistencies in MSIS enrollment data, where possible.
- It is linked to the Medicare Enrollment Database (EDB) to improve identification of Medicare-Medicaid dual enrollees. Further, some additional information from the EDB—of potential interest to researchers—is added to MAX enrollment records.

- It links selected eligibility and demographic information for each enrollee to each claim and links a summary of use and expenditures to each enrollee record.
- It links MAX prescription drug claims to code systems to identify drug therapeutic use classes and groups.[3]

Each year of MAX data consists of a person summary (PS) file and four claims files—inpatient (IP), institutional long-term care (LT or ILTC), prescription drug (RX), and other services (OT)—for each state and the District of Columbia. MAX contains both fee-for-service (FFS) and managed care claims data. For FFS enrollees, detailed data on service use, diagnoses, and cost of care are available. For managed care enrollees, MAX includes capitation premium payments to managed care plans and encounter data for the services provided by those plans. In Exhibit 1, we explain the MAX file types in detail. CMS produces MAX data as soon as all of a state's required quarters of MSIS data are available. Because the files are produced by state, some states may have MAX data available for more years than others. For example, the MAX 2010 data set includes all states, but the 2011 data set currently includes 45 states, and the 2012 data set includes 35 states.[4]

CMS is currently working with states to transition from reporting in MSIS to reporting in Transformed-MSIS (T-MSIS). T-MSIS differs from MSIS in several ways, including (1) a shift from quarterly to monthly reporting, (2) the addition of new file types (third-party liability, provider, and managed care plan data), and (3) the addition of many new data elements and modification of existing data elements. Because the MAX data set is derived from MSIS, the adoption of T-MSIS is expected to result in changes to the MAX files as states begin to replace MSIS submissions with T-MSIS data. The specific timing of the transition to T-MSIS will vary by state.

MAX data resources

Several MAX-related resources are publicly available on the CMS website to assist MAX data users and to support research and policy analysis on Medicaid and CHIP populations using MAX data. CMS produces a new version of each resource with each release of MAX data. In this section, we highlight two primary resources—the validation tables and the anomaly tables. In addition, we briefly identify other MAX data resources.

Exhibit 1. MAX file types

Person summary (PS) contains summary information about each Medicaid enrollee during the year, including demographic and enrollment data as well as service utilization and expenditure measures.

Inpatient (IP) contains records for inpatient hospital stays. Claims records in this file:

- Are bundled into a "stay." A single inpatient visit that generates multiple claims has one record.
- Provide information—using Uniform Billing (UB-2004) revenue codes—on inpatient hospital facility services, including accommodations, supplies, medications, laboratory services, x-rays, surgical suites, and other services provided by a facility while a patient is hospitalized.
- Include multiple diagnosis and procedure codes.
- Generally do not include information on physician and other provider services provided while a patient is hospitalized (other than procedures). The OT file includes these services.

Long-term care (LT or ILTC) contains records for services provided in long-term care institutions. Claims records in this file:

- Are typically weekly or monthly claims for room and board or ancillary services provided by the facility.
- Include diagnosis codes, but these codes may reflect patient diagnoses at admission to the facility.
- Include records for nursing facilities, intermediate care facilities for individuals with intellectual disabilities, inpatient psychiatric services for enrollees under age 21, and services provided at mental hospitals for the aged.

Prescription drug (RX) contains records for prescription drugs and other services provided by free-standing pharmacies. Claims records in this file:

- Include durable medical equipment (DME) or vaccines billed by pharmacy providers that include a national drug code (NDC).
- Do not include the above services provided during inpatient hospital visits or for nursing facility residents in some states.

Other services (OT) includes all types of services that are not included in the IP, LT, or RX files, including physician and other provider, clinic, home health, laboratory, x-ray, dental, outpatient hospital, and prescription drugs administered by a physician or other provider.

Validation tables. For each year of MAX data, a set of validation tables is available for each state, including a table for each of the five MAX file types. The tables include over 1,000 measures (called validation measures) that report enrollment counts, service use, and expenditures for key subpopulations of Medicaid and CHIP enrollees in the state. The validation tables include data from the reporting year and compare them to the two previous years of data. Researchers may use the validation tables to obtain state-level statistics on Medicaid enrollment and expenditures and to assess the accuracy and completeness of MAX reporting on a given topic or subpopulation.

Anomaly tables. The claims and eligibility anomaly tables include key summary statistics from the validation tables for each state to provide high-level overviews of MAX data for a given year.[5] The tables expand on the summary statistics by highlighting data quality issues and unusual patterns in specific states, including context on data quality issues when such information is available. In addition, the tables provide a side-by-side comparison of state data and an overview of key state Medicaid program characteristics. The tables compare selected measures against expected ranges or thresholds for annual change to flag larger-than-expected changes in the data. In general, the tables provide insights into:

- **Data issues.** The tables highlight data inconsistencies or reporting errors that are observed in each state's data and affect the completeness of reporting for data elements.
- **Changes to state programs.** The eligibility tables note substantial changes to state programs that affect key data elements, such as new waivers or managed care programs.
- **Differences with other data sources.** During the validation process, MAX is compared to external benchmarks, including other CMS sources on Medicaid (such as the Statistical Enrollment Data System (SEDS) reporting system for CHIP, the Managed Care Enrollment Reports, the EDB, and the Demonstrations and Waivers website) as well as Administration for Children & Families's data on Temporary Assistance for Needy Families (TANF) enrollment. The eligibility anomaly tables note any discrepancies between MAX and these data sources—and their causes, when known.

Assessing MAX data quality: expected ranges

For many validation measures, CMS has identified an "expected range" of values based on the typical range of values for the measure in most state Medicaid programs historically. The validation and anomaly tables flag values that do not fall within the expected range for a measure within a year or across years. The flags are a key resource for identifying potential trouble spots or idiosyncrasies in a state's Medicaid data. They may also reflect differences in state Medicaid programs. When CMS has additional information about why a measure is out of range in a state, it provides the information in the anomaly tables.

In Exhibit 2 (on page 4), we provide examples of the types of data quality issues that may be identified by using the expected ranges for several data elements in the anomaly tables and the types of information CMS may provide about data quality issues.

Other MAX data resources

- **Data dictionaries.** The MAX data dictionaries provide a detailed, comprehensive list of data elements contained across all MAX files for each year, including source information and user notes. Each MAX file type has its own data dictionary.
- **Waiver crosswalks.** Every state has at least one Medicaid waiver that grants the state flexibility in providing Medicaid services. Most states maintain several waivers of different types. Beginning with MAX 2005, MAX waiver crosswalks identify active Medicaid waivers in each state for the year. The crosswalk for each state includes the following information for each waiver: waiver name, waiver type, waiver identification number (ID), and start and end dates.
- **Managed care crosswalks.** Beginning with MAX 2009, a managed care crosswalk is available by year. The crosswalk identifies each state's managed care plan IDs, plan names, plan types, and whether each plan reported enrollment, capitation payments, and managed care encounter records.
- **Frequently asked questions (FAQs).** FAQs address many basic questions about using MAX data, such as how to identify different populations or construct analyses.

Exhibit 2. Example state data anomalies, MAX 2010

Data element(s)	Anomaly table measure	Expected values	Number of states in expected range	Anomalous example(s)	Reason
Social Security Number (SSN)	Percentage of records with invalid or missing SSN	0.0–5.0	43	California: 34.5%	About two-thirds of these enrollees qualify only for family planning benefits, and California does not collect SSNs for family planning–only enrollees.
Waiver Type and Program Type	Percentage of 1915(c) waiver enrollees with no waiver claim	0.0–15.0	35	Maine: 100% Kansas: 100%	HCBS claims were not available for either state in 2010.
Race	Percentage of enrollees with missing race	0.0–10.0	26	Colorado: 58.0% Iowa: 45.7% Massachusetts: 53.5% Rhode Island: 56.4%	These states do not require race information to be reported as part of the enrollment process.
Procedure Code	Percentage of claims (in OT file) with a procedure code among fee-for-service non-crossover claims[1]	95.0–100.0	37	Arizona: 30.6%	N/A
Inpatient Length of Stay	Average length of stay (in IP file) among fee-for-service non-crossover claims	Values above 2 standard deviations from the arithmetic mean (across states) are considered anomalous	50	Rhode Island: 8.2 days	N/A

Source: Mathematica analysis of MAX anomaly tables, 2010.
N/A = Not available (claims anomaly tables do not generally provide explanations for data quality issues that have been flagged).
HCBS = Home- and community-based services; IP = inpatient; OT = other services.
[1]Non-crossover claims are Medicaid claims for Medicaid only enrollees, and therefore do not include Medicare cost sharing amounts.

- **File specifications.** For users of MAX, some supports are available for data processing, including Statistical Analysis System (SAS®) load statements for each MAX file and a table of record counts to verify the number of records in each MAX file.

Using MAX data for research projects

MAX files contain a wealth of information about the Medicaid program. Given Medicaid's complexity and the variation in programs and data quality across states, researchers should consider a number of factors when selecting states, study populations, and data elements for inclusion in their analyses. For example, in an analysis of hospital admission rates, the inclusion of individuals who receive very limited Medicaid benefits (such as only family planning services) could lead to biased results. In addition, the completeness and quality of data elements vary by state, and some states do not report all elements. Below, we provide an overview of some key data elements that are likely to be of interest to researchers. We include a discussion of tips for using the elements and provide illustrative examples of how they may be used in research studies. Next, we describe specific data issues and approaches to determining which states and data elements are appropriate for use in a given study.

Using key data elements to define research populations and services

Population for analysis. MAX data reflect the diversity of Medicaid enrollees and include all individuals with a Medicaid or M-CHIP enrollment or claim record during the year,[6] from individuals who were enrolled and

used many Medicaid services during a year to individuals who were eligible only for narrowly defined benefits. Researchers should consider the appropriate population for their analyses. In Exhibit 3, we identify key Medicaid populations that may be relevant to researchers and describe how researchers may use the MAX PS file to focus on their selected population.

Assessing service use and expenditures. As discussed above, MAX includes several different claims files. Across all claims files, the Type of Claim Code data element allows users to differentiate between claims to identify services provided through FFS and services provided through capitated managed care. Types of claims include:

- **FFS claims.** These claims report individual services and total expenditures for each service provided on an FFS basis. They are included in all four MAX claims files. Several MAX data elements provide information on payments made to providers for Medicaid enrollees. The Medicaid Payment Amount data element provides information on Medicaid payments for FFS services, including both federal and state contributions. Users should be aware that the Medicaid paid amount for a given service does not equal the total cost of care when an enrollee is dually covered by Medicare, when an enrollee has private insurance, or when an enrollee has any out-of-pocket liability. However, the payment does represent the total *Medicaid* expenditure.

Exhibit 3. Key Medicaid populations for research

Population	Description	Research uses
All individuals with Medicaid/CHIP records	All individuals with MAX records, including individuals who were enrolled in M-CHIP or separate CHIP during the year as well as individuals who had claims records but no associated Medicaid enrollment records. **How to identify:** Include all records in a state's file.	Provides a complete picture of total Medicaid as well as M-CHIP expenditures and coverage. Enrollment information for separate CHIP is provided for some states.
Medicaid/CHIP enrollees	Individuals who were enrolled at some point during the year. **How to identify:** Include records for individuals with at least one month of Medicaid or CHIP enrollment. The monthly CHIP Code data element may be used to distinguish Medicaid, M-CHIP, and separate CHIP enrollees.	Limits analyses to individuals identified by the state as program enrollees. As noted above, not all states provide separate CHIP enrollment information.
Full-benefit Medicaid/M-CHIP enrollees	Individuals who received full (or equivalent) benefits during the year. **How to identify:** Include Medicaid/M-CHIP enrollees based on monthly codes in the Restricted Benefit Flag data element. Commonly, full- or equivalent-benefit groups exclude individuals who are eligible only for family planning, emergency services, Medicare cost sharing, and private insurance premium assistance. Conversely, researchers may want to use this data element to focus on a specific population with restricted benefits.	Limits analyses to individuals who are eligible for full benefits or broadly comparable benefits; enrollment patterns and expenditures may differ between individuals who receive full benefits and individuals who receive only limited services.
Fee-for-service Medicaid/M-CHIP enrollees	Full-benefit Medicaid/M-CHIP enrollees who receive most or all of their services on a fee-for-service (FFS) basis. **How to identify:** **1.** To identify individuals who receive most or all of their services on an FFS basis, include full-benefit Medicaid/M-CHIP enrollees who are not enrolled in any of the following types of capitated managed care plans during the year:[1] comprehensive managed care (health maintenance organizations (HMOs), Health Insuring Organizations (HIOs), or Programs of All-Inclusive Care for the Elderly (PACE)), behavioral health plans, and long-term care plans. **2.** To identify individuals who receive all of their services on an FFS basis, exclude individuals with any capitated managed care enrollment, including enrollment in dental plans. Include people enrolled in primary care case management programs (PCCMs).[1]	Limits analyses to the population that is generally expected to have the most complete service use and expenditure data in MAX. Because MAX contains limited service use data for enrollees covered by managed care, researchers may want to focus on individuals who have primarily FFS claims.[2]

continued

Exhibit 3. Key Medicaid populations for research *(continued)*

Population	Description	Research uses
Medicaid/M-CHIP managed care enrollees	Medicaid/M-CHIP enrollees who are enrolled in managed care plans. Medicaid and CHIP enrollees may be served by comprehensive managed care plans, such as HMOs and HIOs, or plans that offer a limited scope of coverage, such as dental or behavioral health services, known as Prepaid Health Plans (PHPs). **How to identify:** To analyze the population ever enrolled in capitated managed care during the year, include Medicaid/M-CHIP enrollees who are identified as enrolled in one or more capitated managed care plans for one or more months. To analyze only the population enrolled in capitated managed care for the entire year, include Medicaid enrollees who are identified as enrolled in one or more capitated managed care plans for all 12 months.	Allows for analyses of the managed care delivery system, such as determining the average number of managed care plans in which enrollees are concurrently enrolled. Also allows for analyses of specific types of capitated managed care plans or outcomes for individuals enrolled in capitated managed care. Over half of all Medicaid enrollees are enrolled in comprehensive managed care plans.
Medicare-Medicaid dual enrollees	Medicaid enrollees who are also enrolled in Medicare, including individuals who receive full Medicaid benefits as well as individuals whose Medicaid benefits are limited to Medicare cost sharing. **How to identify:** The monthly Dual Code data element identifies Medicare-Medicaid dual enrollees and indicates whether an enrollee was eligible for full or restricted Medicaid benefits during each month. Conversely, because Medicare-Medicaid dual enrollees often have unique service use and expenditure patterns, researchers may want to limit their studies to individuals who are not dually enrolled in Medicare.	Medicare-Medicaid dual enrollees are a unique Medicaid population. Their inclusion/exclusion from a study will likely affect service use and expenditure results,[3] particularly results for aged individuals or individuals with disabilities.
Eligibility categories	Eligibility category groups include a variety of subpopulations of Medicaid/M-CHIP enrollees based on their pathways to Medicaid/M-CHIP eligibility and/or their demographic characteristics. • The pathway-to-eligibility categories identify how each enrollee became eligible for Medicaid/M-CHIP; the categories include cash assistance enrollees, medically needy enrollees, poverty-related enrollees, enrollees eligible via Section 1115 expansion waivers, and other pathways. • Demographic groups include enrollees who are: aged, have disabilities, children, and adults without disabilities. Enrollees in the disability category may be of any age. **How to identify:** The monthly Uniform Eligibility Group (UEG) data element has two digits. The first digit identifies an enrollee's pathway to eligibility, and the second digit identifies the enrollee's demographic category.	Allows researchers to limit analyses to specific subpopulations of interest.[4]

Source: Mathematica analysis of MAX data.
M-CHIP = Medicaid-expansion CHIP.
[1] As discussed below, capitated managed care plans are plans that receive a set monthly premium payment to provide all covered health services to enrollees. PCCMs are not capitated. Under PCCMs, services are paid on an FFS basis, and primary care providers are paid an additional small monthly fee for conducting care management.
[2] Service use data (encounter records) in MAX for managed care enrollees are of varying quality and completeness across states and do not include Medicaid payment amounts.
[3] Medicare pays for Parts A-, B-, and D- covered services before Medicaid, and MAX includes only claims information for payments made by Medicaid. These may be copayment and/or deductible amounts up to payment limits established by each state. For that reason, estimates of service use and expenditures for Medicare-Medicaid dual enrollees often differ from those of individuals who are not enrolled in Medicare.
[4] Researchers should be aware that states may vary in their reporting of these groups. For example, some states include all enrollees with disabilities in the disability category while others include enrollees with disabilities over age 65 in the aged category.

- **Capitation claims.** These claims report premium payments to managed care plans for risk-based capitated coverage. They do not identify services provided under capitated coverage. They are typically reported on a monthly basis per enrollee. Capitation claims are included in the OT claims file. They include managed care Plan IDs so that researchers may identify the managed care plan and type of plan (including comprehensive managed care plans and prepaid health plans, such as behavioral, dental, prenatal, long-term care, and other plans). In MAX, monthly payments to providers to provide case management services for primary care case management programs (PCCMs) are reported as capitation claims. However, PCCM payments are not risk-based and should not be equated with capitation payments for risk-based plans. To calculate the total amount Medicaid paid for the services provided to an enrollee through a capitated managed care plan, users may sum the total capitation payments made for that beneficiary to that plan.
- **Encounter records.** These records report the services that a capitated managed care plan provided to an enrollee. The direct cost to Medicaid of a specific service under capitated managed care does not apply because the plan is paid a fixed amount, typically per month, to provide a defined set of services. Encounter data reflect an amount paid (by Medicaid) of zero.[7]

Researchers who are interested in analysis at the delivery system level may want to focus on a specific type of claim. For example, someone analyzing services provided via capitated managed care should focus on encounter records, whereas someone interested in analyzing services or expenditures provided on an FFS basis should focus on FFS claims. In contrast, researchers interested in person-level analysis will need to identify all types of claims for each enrollee to develop a total picture of service use and expenditures for that individual.

Assessing MAX data quality and completeness

Common data quality concerns across states. There are some MAX data elements that have historically been less reliable or complete across states. When considering use of these data elements, researchers should closely examine the anomaly tables for state-specific notes on the completeness and accuracy of these fields in their selected states.

- **Race/ethnicity.** About half of states report 10 or more percent of their enrollees with a missing race field. Among these, eight states report missing race for 40 or more percent of their enrollees in 2010. In some cases, this occurs because race is not a required element in a state's Medicaid data system. When race/ethnicity is reported, accuracy may vary across states.
- **TANF.** Approximately two-thirds of states do not provide any information about TANF enrollment and 9-fill this data element.
- **Waiver enrollment.** In some cases, states are unable to report enrollment in certain waivers or waiver types in MAX. In many cases, the issue occurs in the months after a waiver is first implemented. The waiver cross-walk provides information about active waivers that are not reported in MAX.
- **Payment amounts for private insurance (third-party payment) and Medicare coinsurance and cost sharing.** These data elements are not widely or completely reported; researchers should be cautious when using them.
- **Procedure coding in claims.** Diagnosis codes reported in MAX data follow the standard national coding system: the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). However, in early years of MAX data, there is considerable variation between states in terms of the procedure coding schemes for Medicaid claims. Claims in the IP and OT files have a mix of types of procedure codes. For IP procedures, ICD-9-CM codes are almost always used. For OT medical services, the codes used are often from the Current Procedural Terminology, 4th Edition (CPT-4)/Healthcare Common Procedure Coding System (HCPCS) Level 1 (national). Many states use local codes for certain services, including home- and community-based services (HCBS); however, MAX includes an HCBS Taxonomy Code data element to standardize the categorization of national and state-specific codes across states for Section 1915(c) waiver services. In addition, some states (such as New York and Ohio)

have used state-specific codes for other types of claims. Data users should determine the coding systems in use in the states and time periods they plan to analyze and be aware that statistical programs that rely on national coding systems may not work with claims for states with high use of state-specific codes. Users may identify the coding systems used by a given state by looking at the Procedure Coding System data element. If the codes used are not standard national codes, the data element will have a value of "other."

- **New data elements.** Researchers may find that new data elements vary in quality when they are introduced but may improve over time. For example, in the claims files, the provider taxonomy and National Provider Identifier (NPI) data elements exhibited issues with completeness, quality, and consistency when they were first introduced but have since improved.

State-specific data issues. Limitations or changes in states' data reporting systems or idiosyncrasies in state Medicaid or CHIP programs may result in state-specific issues related to quality or completeness of data elements. Researchers should note that the quality or completeness of a state's data may change over time and that changes may be limited to specific data elements. As discussed above, CMS documents known data quality issues in the MAX anomaly tables, and researchers should consult these resources to learn more about data issues that could affect the usability of a state's data to answer a specific research question for a specific time period. Below, we describe some common types of state-specific data issues that CMS has identified in MAX data.

State-specific data quality issues in MAX take many forms. Concerns about the quality of a state's MAX data may be limited to a few data elements or may universally affect a state's data. Similarly, data issues may either occur for a short period following a program or data system change or persist for multiple years. One common cause of data quality changes in a state is an overhaul or replacement of the state's data reporting system (often referred to as its Medicaid Management Information System—MMIS). MMIS updates often result in changes in many MAX data elements that may translate into either improved reporting or new data quality concerns. Researchers should consult the anomaly tables for an indication of whether changes are believed to be errors. In some cases, CMS may flag a known error or omission that resulted from a change in reporting. In other cases, MAX documentation may note a marked change in a data element, although it may be unclear which data are more accurate—the old or new data. Researchers should also note that states may undertake efforts to correct or improve reporting over time. For example, following targeted technical assistance by CMS that started in 2010, several states improved the quality and completeness of their reporting for enrollees in separate CHIP programs as well as for encounter data records. In Exhibit 4 (on page 9), we include examples of the range of state-specific data quality concerns that CMS has documented in recent years of MAX data.

When assessing data quality in a state, researchers should keep in mind that, although MAX data resources provide general insights into how most state programs report data and explain known data reporting anomalies, state data may differ from one state to another for reasons that do not necessarily reflect data quality issues. For example, under federal guidelines, states have considerable flexibility in determining their Medicaid programs' eligibility criteria and benefits. In addition, state demographics can

Data quality issues: factors to consider

- **Prevalence.** Is it a widespread issue that affects many data elements and populations of enrollees? If limited, it may not affect data quality and reliability for the subpopulation of interest.
- **Duration.** Is it an issue that appears to affect multiple months or years of data? In the case of using multiple years of MAX data, will it affect comparability within a state over time? For a one-year analysis, lack of comparability may not be as critical as for an analysis extending beyond one year of data. However, a major change in coding and reporting could be critical to a trend analysis even if the new data are not inherently incorrect.
- **Severity.** Is it clear from the MAX data documentation whether the data concern is an issue of unreliable data or a case of data changes that led to the data becoming less comparable to data in another time period?

Exhibit 4. Examples of state-specific data changes or quality concerns in MAX data, 2010–2012, by characteristics

Type of data issue	State: description	MAX years affected
Isolated issue	Oregon: Family planning enrollees are not reported in MAX.	1999 and ongoing
Isolated issue	Maine: Implementation of new MMIS resulted in termination of all Medicaid waiver enrollment reporting, although the state's waivers are still active.	October 2011 and ongoing
Isolated issue	Florida: Family planning enrollees are underreported.	2011 (January–June)
Widespread issue	Kansas: All claims files were unavailable so MAX included no service use or expenditure data.	2010
Widespread issue	Maine: IP, LT, and OT claims files were unavailable so MAX included no service use or expenditure data for the services reported in these claims files.	2005 and ongoing
Data improvement	Mississippi: The state enhanced MSIS reporting to include separate CHIP enrollees.	2011 and ongoing
Data improvement	Ohio: The state improved the completeness of claims data by adding encounter data records.	2011 and ongoing
Data change	Massachusetts: Coding changes were implemented in a new MMIS in 2009 and resulted in notable enrollment shifts in several key enrollment data elements. It is unclear whether the data changes represent a data quality improvement or error.	2010 and ongoing

Source: Mathematica analysis of MAX anomaly tables, 2005-2012.
IP = inpatient; LT = long-term care; MMIS = Medicaid Management Information System; MSIS = Medicaid Statistical Information System; OT = other services.

affect the composition of a state's Medicaid program, causing enrollment, service use, and expenditures in Medicaid to vary substantially from state to state. Finally, changes to state Medicaid programs, such as a new managed care program, might cause legitimate and expected changes in the population represented in state Medicaid data from one year to the next. When conducting analyses of Medicaid across multiple states, researchers should consider the characteristics of specific state programs in making decisions related to study design and interpretation of findings.

Accessing MAX data and resources

The CMS MAX website provides information and resources for researchers using MAX data. It includes updates on data availability, access to data resources, chartbooks describing Medicaid enrollment and service use patterns, and issue briefs on special topics. The website provides access to the anomaly and validation tables as well as to other resources described in this issue brief. CMS updates the website with new state data as the data become available. The home page is at http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidData-SourcesGenInfo/MAXGeneralInformation.html and access to other MAX pages is available there.

To access complete, person-level MAX data, researchers must enter into a data-use agreement with CMS. Further information about how to obtain the data is available at the Research Data Assistance Center (ResDAC) website at http://www.resdac.org/cms-data.

Endnotes

[1] The CMS MAX website provides several examples of research conducted with the MAX data system. Further information about the website appears at the end of this issue brief.

[2] MAX data do not include information about Medicaid or CHIP enrollees in Puerto Rico or other U.S. territories. For enrollees in separate CHIP programs, the data set contains only limited enrollment information and no service utilization or expenditure information.

[3] Access to the prescription drug data is limited to researchers covered under a CMS licensing agreement.

[4] CMS releases new state data periodically. Researchers should consult the CMS MAX website to learn which states' data are currently available. Further information about the website appears at the end of this issue brief.

[5] Starting with MAX 2007, MAX data anomalies are provided in a table format to give researchers an easier reference to compare and contrast data across states. Early versions of the anomalies through MAX 2007 were produced as state-specific narrative reports. In MAX 2007, anomalies were produced in both text and table formats. In MAX 2008 and thereafter, anomalies are produced in table format only.

[6] Separate CHIP enrollment information is also available in some states.

[7] States are directed to record the amount that an MCO paid to providers in the Amount Charged field. However, the data element is not well reported or validated.

For further information on this issue brief series, visit our website at www.mathematica-mpr.com

Princeton, NJ • Ann Arbor, MI • Cambridge, MA • Chicago, IL • Oakland, CA • Washington, DC