MEDICAID 1115 DEMONSTRATIONS

Working with Medicaid Managed Long Term Services and Supports (MLTSS) Data: Challenges and Solutions for State and National Evaluators

Jenna Libersky, Su Liu, Madeline Pearse, and Debra Lipson

Introduction

Over the past decade, state Medicaid agencies have increasingly shifted the delivery of long-term services and supports (LTSS) for older adults and people with disabilities away from fee-for-service (FFS) systems, which pay providers for each service delivered. Agencies have instead shifted these services to managed care delivery models in which they contract with private managed care plans to arrange and pay for LTSS. As of July 2016, 20 states offered 24 managed LTSS (MLTSS) programs, enrolling more than 1 million users (Centers for Medicare & Medicaid Services [CMS] and Mathematica Policy Research 2018). This is a significant increase since 2004, when only 8 states had MLTSS programs enrolling about 100,000 users (Saucier et al. 2012).

MLTSS programs are intended to accelerate the shift from institutional to home and community-based services (HCBS), improve service delivery and quality of life for beneficiaries, and save costs or limit cost growth. Many states have turned to MLTSS because they believe the model has a greater ability than traditional FFS models to achieve these goals. Consequently, it is important to monitor progress toward these goals and evaluate the success of MLTSS relative to FFS in achieving expected results. Meaningful monitoring and evaluation require high quality data; any problems in the completeness, accuracy, consistency, or timeliness of the data may produce unreliable or misleading findings.

About this brief. This brief is designed to help state and national evaluators identify common challenges in using and analyzing MLTSS data, understand the limitations of these data, and minimize the risks of using faulty data. The brief draws on interviews with Medicaid officials involved in evaluating and monitoring MLTSS programs; external evaluators of MLTSS programs; and Mathematica staff who have experience using MLTSS data in plan, state, and national data sets. It also draws on Mathematica's experience assessing MLTSS programs in two states for the interim MLTSS evaluation report to CMS, part of the national evaluation of section 1115 demonstrations (see text box on page 10). (For details on the data collected, see the Methods and Data Sources text box on the next page.)

THE MEDICAID CONTEXT

Medicaid is a health insurance program that serves low-income children, adults, individuals with disabilities, and seniors. Medicaid is administered by states and is jointly funded by states and the federal government. Within a framework established by federal statutes, regulations and guidance, states can choose how to design aspects of their Medicaid programs, such as benefit packages and provider reimbursement. Federal law also specifically authorizes experimentation by state Medicaid programs through section 1115 of the Social Security Act. Under section 1115 provisions, states may apply for federal permission to implement and test new approaches to administering Medicaid programs that depart from existing federal rules yet are consistent with the overall goals of the program and are budget neutral to the federal government.

For the past two decades, states have increasingly turned to private managed care plans to deliver long-term services and supports (LTSS) to Medicaid beneficiaries with disabilities. Section 1115 is one of several federal authorities that states can use to operate managed long-term services and supports (MLTSS) programs. In contrast to fee-for-service, which pays providers for each service they deliver, states that operate MLTSS programs pay managed care plans a fixed per-member-per-month (PMPM) amount to provide all covered services for enrollees. The capitated PMPM payment arrangement – combined with contract requirements to protect enrollees – can create an incentive for the plans to improve care coordination, reduce unnecessary services, and increase the availability of home and community-based alternatives to institutional care.

In February and March of 2018, we conducted eight 30- to 60-minute semistructured telephone interviews with Medicaid officials involved in evaluating and monitoring MLTSS programs; external evaluators of MLTSS programs; and Mathematica staff who have experience using MLTSS data in plan, state, and national data sets. Interviewees shared their experiences with each type of data commonly used to monitor and evaluate MLTSS programs, along with the challenges they encountered obtaining high quality data. We also asked about techniques to identify and remedy issues found in data. We drew additional information from the data limitations reported in 1115 annual reports and evaluations on Medicaid.gov. Interviewees were given an opportunity to review a draft of this brief for accuracy.

The brief begins by presenting common questions that MLTSS monitoring and evaluation efforts would typically answer, and then describes the data sources that inform each question. It proposes a framework for high quality data and identifies known challenges to the quality of each data source. It also describes validation techniques that evaluators can use to identify data issues. Finally, it discusses approaches that can mitigate the risks of using faulty data for monitoring and evaluation. Though monitoring and evaluation differ in their aims and methods (Exhibit 1), both activities rely on high quality data, often from the same source.

This brief describes MLTSS data challenges in general terms so that program administrators and evaluators at state or national levels can apply the findings to their work. Although state and national data sources differ in some respects, they are often connected. Data problems that originate with managed care plans get passed on to the state and, if not corrected, passed on to databases maintained by CMS (for example, the Transformed Medicaid Statistical Information System, or T-MSIS). Unless otherwise specified, problems with a particular data set—and the validation techniques used to detect the problems—apply to data at both state and national levels.

Exhibit 1. Comparing monitoring and evaluation activities

Monitoring and evaluation are two distinct activities that inform our understanding of the implementation and outcomes of MLTSS. The table below distinguishes monitoring from evaluation based on what is being measured, how often and over what time period measurement is conducted, who conducts the measurement, what data are required, and what methods are used.

Distinguishing feature	Monitoring	Evaluation
What is being measured?	Monitoring focuses on the outputs of implementation, comparing what is delivered to what was planned. It might answer questions such as: how many people are enrolled; how many services of each type are delivered; how many people are assessed within the time frame specified in the contract; and how close are these numbers to expectations?	Evaluation focuses on the effects of the program and whether the program achieved the intended outcomes. It also measures secondary effects (intended or not) and looks for lessons learned from the program that can be applied elsewhere. It answers questions such as: how substantial and valuable were the outcomes; how well did they meet enrollees' needs and goals for care; and were the impacts practically significant?
How often and over what time period is measurement conducted?	Monitoring is an ongoing function that takes place at regular intervals throughout program implementation. The intervals are often shorter than what is examined through evaluation.	Evaluation is a discrete activity that involves assessing the entire program cycle, often retrospectively and over longer time periods than monitoring.
Who conducts the measurement?	Monitoring is usually done by people who are directly implementing or overseeing the program.	Evaluation is best conducted by an independent outsider whose position allows for an impartial review.
What methods are used?	Monitoring questions are not fixed and can change as a program evolves. Measures often involve simple counts or ratios that reveal trends in program data relative to general expectations.	Evaluation answers a predetermined set of questions. Relative to monitoring, it involves the use of more sophisticated methods to compare program trends to an alternative (that is, a comparison group). Evaluators use statistical controls to adjust for population variation (for example, differences in the risk of needing services within an enrolled population).
How does monitoring relate to evaluation, and vice versa?	The data collected and insights gained through monitoring are often fed into the evaluation process.	Evaluation draws on data collected and used in monitoring, but it also may involve use of additional data and measurement techniques.

Source: Adapted from Tatian, Peter. "Performance Measurement to Evaluation." Washington, DC: Urban Institute and the World Bank Group Community Outreach Program, March 2016.

MLTSS evaluations typically seek to answer questions on enrollment, utilization or realized access, spending, rebalancing LTSS in favor of HCBS care, enrollee experience, and quality of life (Table 1). The data sources used to answer these questions include the following:

- **Enrollment records**, which contain information on a person's Medicaid eligibility, managed care enrollment, and demographic characteristics.

- **Claims or encounters**, which contain information on each service rendered, including the type of service or procedure, duration, location, provider type, and amount paid. FFS claims include the amount the state paid each provider for the service; the sum of claims for program beneficiaries represents the program's total cost. Managed care encounter records, which mirror FFS claims, often exclude information on payments by the managed care plans to each provider, unless a state requires encounter records to include this information.[1] If provider payments are missing from many encounter records, one can derive the total cost of a program from the sum of capitation payments to managed care plans.

Table 1. Examples of MLTSS monitoring and evaluation questions and data required to answer them

Domain: Sample questions	Enrollment records	Claims or encounters	Functional assessment information	Plan financial reports	MLTSS payment rates from actuarial certifications	Medical records or chart abstractions	Enrollee and beneficiary surveys	Appeals and grievances
Total enrollment: How many people are enrolled in MLTSS plans at any given time?	X							
Enrollment mix: Who is enrolled in MLTSS (what are the demographic, health, functional, or administrative eligibility characteristics of MLTSS enrollees)? How has MLTSS enrollment changed over time—overall and within various demographic, health, functional, or administrative groups?	X	X	X					
Utilization/realized access: How does use of medical care and LTSS compare over time or between MLTSS and any alternative delivery system available to MLTSS enrollees (such as a Financial Alignment Initiative demonstration, FFS LTSS, or Programs of All-include Care for the Elderly [PACE])?	X	X	X					
Total spending: How does total MLTSS spending – in aggregate, per population, or per enrollee – change over time? How does total MLTSS spending compare to spending in an alternative delivery system available to MLTSS enrollees?		X		X	X			
Spending by setting of care: How does the portion of total MLTSS spending on institutional care change relative to HCBS?	X	X				X		
Rebalancing/LTSS setting: How does the use of institutional services compare with the use of HCBS? How many new MLTSS enrollees use HCBS versus institutional care? How many existing institutional residents (short or long term) successfully transition to the community? How many community-dwelling HCBS users are admitted to institutions?	X	X						
Quality of care: How does the quality of care provided under MLTSS change over time or compare with the care provided under FFS or an alternative system (measured through avoidance of falls, wounds or urinary tract infections; vaccination rates; or adherence to medication regimens, for example)?	X	X	X			X	X	X
Needs assessment and care provision: How timely, comprehensive, and person-centered are the assessments of and plans to address enrollees' needs, as well as the care they receive under MLTSS?	X	X	X			X	X	
Access to care: How do enrollees perceive access and quality of care provided through MLTSS? How satisfied are they with their care providers and coordinators? Do patterns of care use suggest that provider networks are adequate?	X	X					X	X
Quality of life: How satisfied are MLTSS enrollees with where they live? Do they feel they have a choice in their daily activities? Do they feel connected with their family, friends, and community? How many who wish to hold a job or volunteer in the community are able to do so?							X	

Unlike FFS claims that providers submit directly to the state, managed care plans are not paid in exchange for the encounter records they submit. However, federal regulations (42 CFR 438.242) require states to collect encounter data from plans and to include such requirements in contracts with plans. Many states also apply penalties or withholds if plans do not submit timely and accurate encounter data files.

- **Financial reports from managed care plans.** As required by 42 CFR 438.3(m), managed care plans submit audited financial reports to the state every year. The information in these reports can help identify the total cost of an MLTSS program; however, the reports do not provide details on the costs of specific populations, services, or settings.

- **MLTSS payment rates from actuarial certifications.** When actuaries develop MLTSS rates for state Medicaid agencies, they prepare detailed information on the monthly capitation amounts to be paid for specific populations. Evaluators can use the plan-specific rates or rate ranges to estimate the cost of MLTSS for certain populations, services, or settings.

- **Functional assessment information,** which describes a person's functional limitations—how much help he or she needs to perform activities of daily living such as bathing, toileting, eating, dressing, and transferring—as they relate to LTSS eligibility. States obtain these data by asking about a person's need for help with daily activities when determining his or her eligibility for LTSS, either in institutions or the community. They may also obtain additional information directly from managed care plans that conduct comprehensive needs assessments with enrollees. The variation in the assessment tools used by states and managed care plans makes it difficult to compare functional needs and thresholds required to receive LTSS across states (Exhibit 2). Because such variation creates a barrier to conducting cross-state

evaluations—and it is beyond the scope of this brief to examine the quality of each state's functional assessment data—this brief does not examine this type of data in detail.

- **Medical records or chart abstractions.** Some quality-of-care measures—including the Healthcare Effectiveness Data and Information Set (HEDIS; National Committee for Quality Assurance [NCQA] 2018) and new measures specific to the MLTSS process of care—require details from medical records or charts to supplement claims or encounters. For example, Mathematica and the NCQA developed and tested four measures of comprehensive assessment and care planning for MLTSS enrollees (CMS, Center for Medicaid and CHIP Services 2018). To construct these measures, trained abstractors were required to review a sample of records or charts. The measures give states and evaluators the ability to use standardized metrics to monitor and compare the quality of care provided by plans.

- **Enrollee and beneficiary surveys,** such as the Consumer Assessment of Healthcare Providers and Systems (CAHPS) for Home and Community-Based Services or the National Core Indicators—Aging and Disabilities, provide information on people's experience of care (CMS 2018; National Association of States United for Aging and Disabilities [NASUAD] 2018). Such surveys are typically conducted with a sample of enrollees at a single point in time or on a recurring basis (for example, every other year).

- **Appeals and grievances.** Federal regulations require managed care plans to have systems in place through which enrollees can appeal adverse benefit determinations or file grievances (for example, regarding dissatisfaction with quality of care; 42 CFR 438.242). States must also review appeals and grievances as part of their ongoing monitoring procedures (42 CFR 438.416), although it is unclear how many states do this in a systematic way.

Exhibit 2. Variation in functional assessment data across states

Federal regulations do not require states to use specific functional assessment tools for MLTSS; therefore, the tools and the data collected vary significantly across and within states—and sometimes within programs themselves (Lewis et al. 2018). As of 2016, at least 124 distinct assessment tools were in use nationwide, with each state using an average of 3 tools, each of which covered different populations (Medicaid and CHIP Payment and Access Commission [MACPAC] 2016). Some states have uniform assessment tools that collect standard information across all MLTSS populations. However, such tools are typically developed for the specific state or program using them (Ingram et al. 2013); only a few programs use national assessment instruments such as InterRAI (Lewis et al. 2018). The variation across tools makes it nearly impossible to compare MLTSS enrollees' level of function across states.

One tool—the CMS-required **Minimum Data Set**—collects quarterly clinical and functional data on all nursing home residents nationwide, regardless of payer (CMS 2017). This rich data set is used to calculate quality measures and can be used to compute length of stay, transitions to the community, and readmissions to facilities for a subset of MLTSS users. The Minimum Data Set offers a couple of advantages over claims: one interviewee reported that data from the Minimum Data Set are 99 percent complete and have a shorter lag time than claims data. Its disadvantage is that it does not provide information on residents of other types of long-term care institutions or on HCBS users in MLTSS, who make up the majority of enrollees.

Defining High Quality Data

To produce reliable results, monitoring and evaluation staff must use high quality data, which are defined by four characteristics (Figure 1):

1. **Completeness:** Data are complete when they have all required information (that is, nothing is missing).

2. **Accuracy:** Data are accurate when they reflect the people enrolled in each plan, the demographic and functional characteristics match those that qualify beneficiaries to enroll in MLTSS and receive LTSS, the services delivered conform to those covered by MLTSS programs, and the volume of service use is within expected parameters.

3. **Consistency:** Data are consistent when the data from each managed care plan or survey vendor are recorded according to the same specification, format, and coding.

4. **Timeliness:** Data are timely when they are relatively current or submitted in accordance with the required schedule.

Figure 1. Four characteristics of high quality data

| Completeness | Accuracy |
| Consistency | Timeliness |
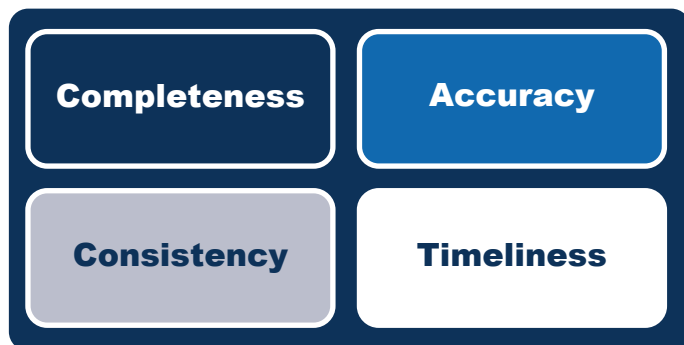
Common Challenges to MLTSS Data Quality

Federal regulations require that states validate much of the data required for monitoring or evaluating MLTSS programs, including encounters and financial reports (see 42 CFR 438.242[d] and 438.602[e]). Validating completeness, accuracy, and consistency requires providing clear guidance to managed care plans on the definition of each data element, service codes, file formats, and other reporting requirements (Byrd et al. 2013). However, the detail of the guidance provided by states varies, as does the rigor of their process for reviewing and ensuring data quality.

When plans receive unclear guidance or do not face any consequences for reporting incomplete or inaccurate data—and when states do not work with plans to address data problems—evaluators are more likely to see incompleteness, inaccuracy, or

inconsistency across one or more data sources. For example, ethnicity and race may be mixed up or missing. Plans may vary in how they report continuing enrollment, with some plans using arbitrary end dates and others leaving the fields blank or filling in placeholder digits (9, for example). Plans may also use different procedure codes for the same category of service. These examples underscore the need for evaluators to work closely with state staff who collect and review MLTSS data to understand its nuances (see the section on Identifying Problems with MLTSS Data Quality).

Evaluators who find that data at the state or national level are insufficient may be tempted to gather data directly from the plans, but there are challenges to doing so. MLTSS plans that are owned by national parent organizations may have national systems and guidelines to follow for reporting encounter data; these systems might not be modified to reflect state-specific coding or reporting requirements, resulting in inconsistent data across plans. Data maintained by the plans may also lack key fields—such as state-assigned categories of services that link to eligibility category, waiver enrollment, or other administrative details—making it difficult to analyze trends across plans.

Evaluators should also keep in mind that plans experience a learning curve for data reporting. Data from early years may not fit established definitions or expectations, especially if states relax the data definitions for plans that have trouble collecting data in order to prioritize timeliness over completeness, accuracy, and consistency. Consistency in data or fields that are submitted voluntarily, especially in the early years of an MLTSS program, will likely vary across plans, reflecting differences in plans' capacity and incentives to report.

Below, we identify several common challenges to MLTSS data quality that can produce unreliable or misleading results. This information is based on Mathematica's history of evaluating state MLTSS programs, including the MLTSS Interim Outcomes Evaluation, and the experiences of the state Medicaid staff and researchers interviewed for this brief.

Enrollment records. One major challenge with enrollment records is that there is only one roster, so there is rarely a source of data against which one can validate enrollment. This makes it difficult to detect inaccuracies or missing fields; as one interviewee said, "You don't know what you don't know." The interviewees were generally confident about the quality of enrollment data maintained by states and submitted to CMS via the T-MSIS or the Medicaid managed care enrollment report; however, they pointed out that many people are required to update fields in the enrollment records, which can lead to inaccuracies. For example, some values (such as living arrangement) are entered by case workers, who may not update the information promptly when someone moves, dies, or becomes ineligible for Medicaid.

Encounter records. LTSS encounter data are subject to some of the same quality challenges as medical services covered under managed care. For example, in states that require encounter records to specify the amount a managed care plan pays to providers, the amounts recorded can vary widely (one state reported seeing one plan pay $100 and another plan pay $1,000 for the same service). One would expect some variation in prices for the same service because plans negotiate with providers on the price, but without additional information, it can be difficult to assess whether the variation is reasonable. In addition, LTSS encounter data may contain services that do not have established procedure codes, such as consumer-directed care that is covered via monthly lump-sum payments to enrollees. Plans may use different procedure codes to record encounters for these non-standard services. The net effect of these challenges is that LTSS encounter records may be less complete and accurate than medical care encounter data.

Timeliness of encounter data is also an issue. Just as FFS claims often have a long run-out period to account for corrections and adjustments, timelines for finalizing encounter data may stretch past the contract-required deadlines. As noted above, timeliness may also be affected if a state does not impose penalties for tardiness or does not use the data to develop capitated rates for the upcoming year. Evaluators must work with states to establish a reasonable run-out time, striking a balance between having complete information in hand and producing timely analyses.

Personal care services are especially difficult to capture accurately in encounter records. In a recent review of such services in the Medicaid Analytic eXtract (MAX) file,[2] the Government Accountability Office (GAO) (2017) found that 15 percent of the records reviewed were missing provider identification numbers, and 34 percent lacked information on the quantity of services provided. State-reported dates of service were also inconsistent across states. For example, some states recorded claims on a single day of service, other states recorded claims over at least a month, and still others recorded claims that covered 100 or more days. This variation within and across states makes it difficult for evaluators to measure the volume of personal care services delivered.

According to GAO, part of the variation across states stems from their use of different methods to verify that home care aides provided billed services to beneficiaries.[3] For example, some states require beneficiaries to sign timesheets, whereas others use electronic timekeeping systems. Federal oversight requirements also vary by authority, contributing to in-state differences in the way that information is categorized in national data sets (GAO 2017). Because personal care services make up a large portion of all MLTSS delivered nationwide in 2015,[4] problems in the quality of personal care encounters make it hard to assess the degree to which enrollees use personal care

and related HCBS versus nursing facility care—an important indicator of rebalancing in the long-term care system.

Medicare claims and encounters for full-benefit dual eligibles. People who are dually eligible for Medicare and Medicaid represent nearly 70 percent of all LTSS users (Medicaid and CHIP Payment and Access Commission [MACPAC] 2014). Although some states exclude dual eligibles from mandatory enrollment in MLTSS programs, these beneficiaries are likely to make up a large share of MLTSS enrollees. In 2016, nearly 2.7 million dual eligibles were enrolled in comprehensive managed care that included LTSS or limited-benefit MLTSS programs (CMS and Mathematica 2018). For full-benefit dual eligibles, Medicare covers hospital, physician, and short-term skilled nursing facility stays, whereas Medicaid pays for Medicare premiums, deductibles, and other cost-sharing, along with LTSS and other services that Medicare does not cover.[5] Many states turn to MLTSS programs to help coordinate medical care and LTSS, which may reduce the use of avoidable hospitalizations and emergency room visits, improve quality, and lower costs.

Evaluating whether MLTSS programs achieve these goals for dual eligibles requires both Medicare and Medicaid data. Interviewees suggested that variables in the Medicaid data may specify whether a claim is a crossover claim, with a portion paid by Medicare. However, without timely and detailed Medicare claims or encounters to link with state-provided Medicaid data,[6] evaluators cannot calculate key measures of health care utilization, such as avoidable hospitalizations, readmissions, or total spending for dual eligibles.

Financial reports or MLTSS payment rates from actuarial certifications. Financial reports document the monthly capitation payments made to MLTSS plans, and rate-setting documentation contains details about how capitation rates are set, sometimes including estimated unit prices for MLTSS-covered services. In theory, evaluators could combine the capitation rates for MLTSS subgroups,[7] as described in actuarial certifications, with the enrollment data on these subgroups to estimate changes in their costs over time. However, interviewees do not use this approach and instead use the financial reports that plans submit to the state primarily to confirm the payments documented in the capitation claims. They did not identify specific challenges to the quality of either data source at the state level, but previous reports have highlighted inconsistencies between financial information available in CMS-64 reports, which states are required to submit to receive federal matching funds, and payment information in the Medicaid Statistical Information System (MSIS) and MAX. For example, GAO (2017) found that personal care expenditure totals from 2012 through 2017 were not reported correctly by states in the CMS-64 reports. Some of these differences occur because the systems that states use to

capture and report adjudicated claims versus actual payments are different (CMS 2018). Evaluators may want to use caution when examining CMS-64 data by service category.

Medical records or chart abstractions. Because plans do not always maintain medical records or charts electronically, information on the process of care may need to be manually extracted from paper charts. This data extraction can be resource intensive, requiring trained personnel to interpret narrative information and validation checks (such as inter-rater reliability studies) to ensure consistency across validators. Information on whether the LTSS care plan developed with the enrollee is shared with the primary care physician—a key indicator of MLTSS coordination—is often not stored in a systematic way, making it difficult to calculate MLTSS process measures. Evaluators who rely on aggregate data reported by the plan, such as indicators of LTSS needs assessments and service delivery, cannot check the data for validity. Plan-reported measures of medical services used by MLTSS enrollees (such as HEDIS) may also combine MLTSS and non-MLTSS users in a single value, making it difficult to draw MLTSS-specific conclusions.

Enrollee and beneficiary surveys that address HCBS or nursing facility care. Though survey data can provide an invaluable window into MLTSS enrollees' experience of care, the way in which surveys are conducted can hurt data quality. First, surveys can be costly, so states may only conduct them once every few years. Even states that administer the same survey over several years may not do so in a way that allows evaluators to see trends or to link survey results to other data sources, such as claims and encounters. Second, the data that surveys produce are subject to recall bias. Finally, low response rates, which are common in surveys of low-income older adults and for surveys such as the CAHPS, can also skew results.

Appeals and grievances. The state-specific processes that enrollees use to report appeals and grievances can make it difficult to categorize and interpret data for monitoring and evaluation. Enrollees and providers may file appeals and grievances for many reasons, but there is no single typology for tracking the nature or subject of these appeals and grievances in Medicaid, making it challenging to monitor trends or to make cross-state comparisons.

The number of appeals and grievances can also be skewed by people who file multiple reports. To file appeals and grievances with a managed care plan, members must be able to reach plan staff and make them understand the issue; for oral appeals, plan staff must document and report the issue (and hopefully address it). Changes in the number of appeals and grievances may reflect how well enrollees advocate for themselves, how well the plan informs enrollees about their rights, or how well a plan executes (and reports on) established processes for filing and addressing issues. One interviewee said that state data were so volatile from month to month that it is hard to identify trends.

Identifying Problems with MLTSS Data Quality

Given the myriad challenges just described, evaluators should review the quality of each MLTSS data source to determine whether it is reliable enough to use for monitoring and evaluation. The four characteristics of high quality data can provide a framework for this review:

1. **Completeness:** Evaluators should examine whether information in key fields used to construct indicators and measures is complete, in terms of the average number of records per enrollee or the percentage of enrollees with records of interest (see Exhibit 3 for sample measures of completeness). If evaluators do not plan to use Medicare data, they should examine completeness separately for beneficiaries who are dually eligible for Medicare and Medicaid versus Medicaid-only beneficiaries, recognizing that data on Medicare-paid physician services, hospital care, skilled nursing facility care, and other such services will not be available for dual eligibles.

2. **Accuracy:** Evaluators should examine whether data in key fields are formatted as expected (for example, whether the fields have the correct number of digits; see Exhibit 3 for additional examples). They should also check that the fields have consistent logic. For example, they might examine whether the date of enrollment begins after a new MLTSS plan is in place, whether race and ethnicity values are in the proper field, and whether the same type of procedures fall under the same category of service and specialty codes. Furthermore, evaluators can examine trends in accuracy over time, accounting for expected changes in enrollment, services covered, changes in level of care, and so on. Internal validation, such as validating minimum and maximum values, medians, and averages in a measure over time, can pinpoint outliers in a given year. External validation— comparing certain data fields across different data sources— can further confirm accuracy, as discussed further below.

3. **Consistency:** Evaluators should compare trends in missing data and accuracy across plans, reporting levels (plan level versus the state level), and over time. This should be done before constructing measures of utilization or care quality because, when reviewing measure values, it can be difficult to discern the amount of variation that stems from expected differences in plan performance versus the quality of the data itself.

4. **Timeliness:** Data will always continue to be updated, and evaluators must strike a balance between having complete information in hand and producing timely analyses. Looking at measures of completeness, accuracy, and consistency over time will help evaluators understand the appropriate amount of run-out time to allow for claims. Evaluators should also speak to state program staff to understand how applicable previous years of data may be to the current program.

Exhibit 3. Examples of encounter data checks from the interim MLTSS evaluation report

For the interim evaluation of section 1115 MLTSS demonstrations, Mathematica investigated the completeness and accuracy of HCBS and institutional encounter data in the Medicaid Analytic Extract (MAX) data set. We focused on the fields required to calculate measures of any long-term care (LT), HCBS use, personal care, and hospital use (three variants) (Libersky et al. 2017). These measures, listed below, may provide other state and national evaluators with ideas on checks to perform before conducting their own evaluations:

- Percentage of participants with at least one HCBS or encounter record
- Average number of HCBS or LT encounter records per MLTSS enrollee
- Percentage of HCBS encounter records with a place-of-service code
- Percentage of HCBS encounter records with a diagnosis code greater than three characters
- Percentage of HCBS encounter records with a procedure code
- Percentage of HCBS encounter records with a CPT-4 or Healthcare Common Procedure Coding System (HCPCS) procedure code
- Percentage of HCBS or LT encounter records with a billing provider ID
- Percentage of HCBS encounter records with a national provider identifier
- Percentage of HCBS encounter records with a servicing provider ID
- Percentage of encounter records with an HCBS taxonomy code
- Percentage of LT encounter records with a service begin date
- Percentage of LT encounter records with a service end date
- Percentage of LT encounter records with a patient status code indicating discharge from or transfer between facilities

CMS also maintains a set of validation tables to help researchers gauge the quality of MAX data related to enrollment, inpatient and outpatient care, and prescription drugs. These tables may provide evaluators with additional ideas on checks related to the completeness and quality of enrollment, claims, and encounter records. Validation tables by state and data year (1999–2013) are located at https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAX-Validation-Reports.html.

Comparing data sources to each other is another important level of review. As one interviewee said, the goal is to "connect as many data sources as possible to make sure you're seeing a similar picture." For example, evaluators can compare the number of capitation records to the number of enrollees, or they can compare the sum of capitation payment amounts to the amounts reported in plan financial statements (one interviewee suggested that the two sources should have at least a 90 percent match). Evaluators can compare the number of MLTSS enrollees to the data reported on state websites or to the data in CMS's Medicaid managed care data collection.[8] They can compare beneficiaries participating in a waiver (via enrollment data) to claims for waiver services. For residents of nursing homes, evaluators can calculate measures using data from the Minimum Data Set and compare the values to measures that plans calculate and report using encounters. Alternatively, evaluators can compare trends in data (particularly in enrollment records by setting of care) to figures that the state reports on its website, and in documentation for stakeholder meetings and legislative hearings or committee meetings.

Communication with state staff. To identify additional data challenges—and to pinpoint the source of any errors or inconsistencies found—evaluators should communicate often with state business intelligence staff or Medicaid Management Information System (MMIS) vendors involved in collecting and aggregating plan-level data. States may not keep documentation on known data issues, so evaluators will likely need to tap into staff's institutional knowledge. State staff involved in data collection and validation can describe any front-end editing processes, technical assistance, or inter-rater reliability studies that test how consistently individual abstractors or surveyors perform relative to one another. Understanding the type of technical assistance provided to managed care plans is especially important to comprehending state data requirements and any common challenges encountered as part of this assistance. For more information on state processes used to validate encounter data, see Byrd et al. (2013).

Minimizing the Risks of Faulty Data

After diligently checking and validating the data, evaluators will often find problematic data fields that affect part or all of the study population. Very rarely are evaluators in a position to completely rectify the issues. However, by recognizing the challenges and understanding their implications, evaluators can minimize the risk of using imperfect data.[9] Several options are discussed below.

Clean the data to the extent possible with clear decision rules. Claims and encounter records can be messy, often containing multiple entries for the same person and same service use. Sometimes, these entries will have conflicting information on important variables such as admission and discharge dates. The same inpatient or nursing-home stay can be captured by multiple records, in which case evaluators need to make clear decision rules and use an algorithm to de-duplicate and roll up the records into unique stays, with reasonable begin and end dates. For example, in the MLTSS Interim Outcomes Evaluation, we found some claims suggesting a person was discharged from a hospital without the facility finalizing their discharge and readmitted the same day or the following day. We treated these overlapping claims in which data indicates that a person is still a patient following discharge as a single stay and required claims suggesting multiple stays to be separated by at least one day.

Replace inaccurate or inconsistent data. Internal or external validation checks may lead evaluators to conclude that certain data fields contain inaccurate or inconsistent data. Though it can be tempting to drop people with such data from the analysis, modifying the data may allow evaluators to maintain a larger sample and thus to draw more statistical power from the analysis.

Evaluators should cross-check information to look for patterns that may lead to decision rules. For example, if an MLTSS enrollee has an enrollment date before the date when the managed care plan actually had a contract with the state, the evaluator could check the total count of enrollees per month from another data source, if available. If universally replacing the inaccurate enrollment data with a date after the contract starts could render a close match to the total count of enrollees, this might be a decision rule that could be applied to retain the enrollees in the analysis.

Fields such as those containing demographic data may appear at a single point in time or over several points in time. If the latter, the values should be consistent over time. But they may not be, in which case the evaluator will need to choose whether to use the value from the most recently available data or to use the most frequently shown value and replace the outliers.

Conduct subgroup analysis or drop people with missing data. When cleaning or replacing data is impossible, evaluators can identify the subpopulations with the most

problematic data and then either analyze their results separately from the rest of the population (that is, conduct a subgroup analysis) or drop them from the study, especially if there is a lot of missing data. The decision depends on the size of the subpopulation and the magnitude of the data-quality problem. Beneficiaries who are dually eligible for Medicare and Medicaid often lack complete data on service utilization because access to Medicare data versus Medicaid data involves entirely different data-user agreements and processes. But some states have so many dual-eligibles that dropping them could mean losing half of the study sample. In such cases, conducting subgroup analyses for duals and nonduals might be a better option. Evaluators could compare the results across subgroups and consider additional sensitivity analyses to assess the robustness of the findings when using different inclusion or exclusion criteria.

Be transparent and include caveats in any reports.
Regardless of the approach used to manage imperfect data, evaluators should be clear and transparent about data quality and any additional data manipulation conducted during the analysis. If findings are documented in a report, evaluators should include a section on caveats or limitations to help readers understand the extent to which data-quality problems or the inclusion or exclusion of certain subpopulations might affect the conclusions.

Conclusion

State MLTSS programs are diverse in terms of providers involved, benefits covered, beneficiaries enrolled, and outcomes that the programs hope to achieve. The data needed to monitor and evaluate the cost, quality, and use of the services delivered by MLTSS plans are also diverse, presenting challenges for evaluators. If such challenges are not identified and addressed, they can produce unreliable and misleading results.

The Medicaid agency staff and researchers interviewed for this brief emphasized that a thorough knowledge of the strengths and weaknesses of each MLTSS data source is critical to effective monitoring and to the success of any evaluation. Before working with such data, evaluators should understand what the data can and cannot provide (that is, the values captured in encounter data fields or their completeness may affect potential answers to research questions). Evaluators should connect and compare as many data sources as possible to see whether they tell a consistent story. Data from the early periods of an MLTSS program—or from a plan's initial participation in the program— should receive special scrutiny because there may be a learning curve associated with submitting quality data.

If evaluators find data gaps or inconsistencies, they should work closely with state Medicaid agencies to understand the source of the problems and to develop solutions. States, in their role as

collectors and reviewers of MLTSS-related program data, should be providing consistent direction for plans to follow and technical assistance to plans to clarify expectations. The details of this guidance and interactions with reporting plans can be especially useful in shaping a monitoring or evaluation strategy.

ABOUT THE MEDICAID SECTION 1115 EVALUATION

In 2014, the Center for Medicaid and CHIP Services within the Centers for Medicare & Medicaid Services (CMS) contracted with Mathematica Policy Research and Truven Health Analytics to conduct an independent national evaluation of the implementation and outcomes of Medicaid Section 1115 demonstrations. The purpose of this cross-state evaluation is to help policymakers at the state and federal levels understand the extent to which innovations further the goals of the Medicaid program and to inform CMS's decisions regarding future section 1115 demonstration approvals, renewals, and amendments.

The evaluation focuses on four types of demonstrations: (1) delivery system reform incentive payment (DSRIP) programs, (2) premium assistance, (3) beneficiary engagement and premiums, and (4) managed long-term services and supports. This issue brief is one in a series of short reports based on semiannual tracking and analyses of demonstration implementation and progress. These briefs will inform an interim evaluation report in 2018 and a final evaluation report in 2020.

References

Byrd, Vivian, Jessica Nysenbaum, and Debra Lipson. "Encounter Data Toolkit." Baltimore, MD: CMS, November 2013. Available at https://www.medicaid.gov/medicaid/downloads/medicaid-encounter-data-toolkit.pdf. Accessed March 23, 2018.

CMS. "CMS.gov: Medicaid Data Sources—General Information." Updated March 19, 2014. Available at https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/index.html. Accessed April 13, 2018.

CMS. "CAHPS Home and Community Based Services Survey." Available at: https://www.medicaid.gov/medicaid/quality-of-care/performance-measurement/cahps-hcbs-survey/index.html. Accessed April 24, 2018.

CMS, Center for Medicaid and CHIP Services. "Measures for Medicaid Managed Long-Term Services and Supports Plans: Technical Specifications and Resource Manual." February 2018.

CMS. "Long-Term Care Facility Resident Assessment Instrument 3.0 User's Manual, Version 1.15." Baltimore, MD: CMS, October 2017. Available at: https://downloads.cms.gov/files/1-MDS-30-RAI-Manual-v115R-October-1-2017-R.pdf. Accessed April 24, 2018.

CMS and Mathematica Policy Research. "Medicaid Managed Care Enrollment and Program Characteristics, 2016." Baltimore, MD: CMS, spring 2018.

CMS." MMLEADS Public Use File: A Methodological Overview." Baltimore, MD, April 2017. Available at: https://www.cms.gov/Medicare-Medicaid-Coordination/Medicare-and-Medicaid-Coordination/Medicare-Medicaid-Coordination-Office/Downloads/MMLEADS_PUF_UserGuide_MethodologicalOverview_042017.pdf. Accessed June 28, 2018.

Contreary, Kara, Katharine Bradley, and Sandra Chao. "Best Practices in Causal Inference for Evaluations of Section 1115 Eligibility and Coverage Demonstrations." Washington, DC: Mathematica Policy Research, June 2018.

Eiken, S., K. Sredl, B. Burwell, and R. Woodward. "Medicaid Expenditures for Long-Term Services and Supports (LTSS) in FY 2015." Baltimore, MD: CMS, April 14, 2017.

GAO. "Medicaid Personal Care Services: More Harmonized Program Requirements and Better Data Are Needed." May 2, 2017. Available at https://www.gao.gov/assets/690/684424.pdf. Accessed March 23, 2018.

Ingram, Carolyn, Alice Lind, and Brianna Ensslin. "Uniform Assessment Practices in Medicaid Managed Long-Term Services and Supports Programs." Princeton, NJ: Center for Health Care Strategies, September 2013. Available at https://www.chcs.org/media/Uniform_Assessment_in_MLTSS_9-6-13_FINAL.pdf. Accessed April 13, 2018.

Lewis, Elizabeth, Steve Eiken, Angela Amos, and Paul Saucier. "The Growth of Managed Long-Term Services and Supports Programs: 2017 Update." January 29, 2018. Available at https://www.medicaid.gov/media/3406. Accessed March 30, 2018.

Libersky, Jenna, Su Liu, Laura Ruttner, Anna Collis, Jonathan Gellar, and Carol Irvin. "Managed Long-Term Services and Supports Design Supplement for the Interim Outcomes Evaluation." Washington, DC: Mathematica Policy Research, June 2017. Available at https://www.medicaid.gov/medicaid/downloads/eval-dsgn-mltss.pdf. Accessed March 28, 2018.

MACPAC. "June 2014 Report to the Congress on Medicaid and CHIP." Washington, DC: MACPAC, June 2014. Available at https://www.macpac.gov/wp-content/uploads/2014/06/June-2014-Report-to-Congress-on-Medicaid-and-CHIP.pdf. Accessed April 13, 2018.

MACPAC. "Functional Assessments for Long-Term Services and Supports." Washington, DC: MACPAC, July 2016. Available at https://www.macpac.gov/publication/functional-assessments-for-long-term-services-and-supports/. Accessed April 13, 2018.

NASUAD. "National Core Indicators - Aging and Disabilities." Available at: http://www.nasuad.org/initiatives/national-core-indicators-aging-and-disabilities. Accessed April 24, 2018.

NCQA. "Summary Table of Measures, Product Lines and Changes: HEDIS 2018, Volume 2." Washington, DC: NCQA, 2018. Available at: http://www.ncqa.org/Portals/0/HEDISQM/HEDIS2018/HEDIS%202018%20Measures.pdf?ver=2017-06-28-134644-370. Accessed April 24, 2018.

Reschovsky, James D, Jessica Heeringa, and Maggie Colby. "Selecting the Best Comparison Group and Evaluation Design: A Guidance Document for State 1115 Demonstration Evaluations." Washington, DC: Mathematica Policy Research, June 2018.

Saucier, Paul, Jessica Kasten, Brian Burwell, and Lisa Gold. "The Growth of Managed Long-Term Services and Supports Programs: A 2012 Update." Baltimore, MD: Truven Health Analytics, July 2012. Available at https://www.medicaid.gov/media/3401. Accessed April 13, 2018.

Endnotes

[1] 42 CFR 438.242(b)(3) requires all contracts between a state and a managed care organization, prepaid inpatient health plan, or prepaid ambulatory health plan to provide for the submission of all enrollee encounter data that the state is required to submit to CMS under 438.818. However, as CMS acknowledged in November 14, 2018 proposed rule, "some states and managed care plans have expressed concern about, and been hesitant to submit, certain financial data—namely, the allowed amount and the paid amount. Managed care plans consider this information to be proprietary and inappropriate for public disclosure." (CMS-2408-P; 83 FR 57264).

[2] GAO reviewed FFS claims and encounters in MAX (which is derived from the Medicaid Statistical Information System [MSIS]) among the 35 states that reported any personal care data in 2012. Although GAO's report combines findings related to FFS claims and managed care encounters, our experience evaluating the quality of personal care encounters for the MLTSS Interim Outcomes Evaluation suggests that GAO's findings are relevant to MAX encounters as well as FFS claims.

[3] Section 12006(a) of the 21st Century Cures Act requires that states implement electronic visit verification (EVV) systems for all Medicaid personal care services and home health services that require an in-home visit by a provider by January 1, 2020 and 2023, respectively. As more states implement EVV systems, the methods they use to verify personal care and home health services should become more standardized within and potentially across states.

[4] In FY 2015, states reported 25.5 percent of all managed care expenditures on CMS-64 reports as "personal care" expenditures. States also reported 16.2 percent of total managed care expenditures as "HCBS under managed care authorities" and 18.2 percent as "HCBS under 1915(c) waivers" (Eiken et al. 2017). The HCBS expenditure categories likely include a large portion of claims for personal care services.

[5] For partial-benefit dual enrollees, Medicaid pays Medicare premiums and, depending on household income, either all or a share of Medicare deductibles and cost-sharing. Partial dual eligibles do not qualify for state Medicaid benefits.

[6] Researchers who are interested in using pre-linked Medicare-Medicaid data from 2006 to 2012 can use the Medicare-Medicaid Linked Enrollee Analytic Data Source (MMLEADS), available through ResDAC. However, interviewees caution that, in addition to the time lag, this linked data set does not contain the same depth of information as the beneficiary-level Medicare information that states might obtain directly from CMS or Medicare Advantage Dual Eligible Special Needs Plans in order to link to Medicaid data. Most notably, MMLEADS contains counts of people who used Medicaid managed care or for whom there are Part A and/or Part B MA premium payments, but it does not contain information regarding particular services used or prevalence of conditions for Medicare or Medicaid managed care enrollees (CMS 2017).

[7] Capitation rates for subgroups correspond to different rate cells. 42 CFR 432.2 defines rate cells as "a set of mutually exclusive categories of enrollees that is defined by one or more characteristics for the purpose of determining the capitation rate and making a capitation payment; such characteristics may include age, gender, eligibility category, and region or geographic area."

[8] Annual Medicaid Managed Care Enrollment Reports are available at https://www.medicaid.gov/medicaid/managed-care/enrollment/index.html.

[9] Researchers should also consider all data limitations when designing their evaluations in order to accommodate the imperfect data to the extent possible. For more information on evaluation design issues concerning 1115 demonstrations, see: Reschovsky, Heeringa, and Colby 2018; and Contreary, Bradley, and Chao 2018.