Catalyzing Medicaid Policy Research with T-MSIS Analytic Files (TAF): Learnings from Year 1 of the Medicaid Data Learning Network (MDLN)

May 15, 2023

Authors: Sarah Gordon, PhD, MS, Boston University; Annaliese Johnson, MPP, AcademyHealth; Susan Kennedy, MPP, MSW, AcademyHealth; John McConnell, PhD, MS, MA, Oregon Health & Sciences University; and William Schpero, PhD, MPH, Cornell University.

AcademyHealth

Introduction

The Medicaid program remains relatively understudied despite providing health care coverage to over 92 million people in the United States as of April 2023. This is due, in large part, to its federalist structure and the lack of a cohesive, national administrative claims data infrastructure. The Centers for Medicare and Medicaid Services (CMS) has made significant efforts to enhance the Medicaid Statistical Information System (MSIS), which compiles data from state Medicaid agencies to inform overall program improvements. In 2019, CMS released the latest generation of federal Medicaid claims data, the T-MSIS (Transformed-MSIS) Analytic Files (TAF), to replace the Medicaid Analytic eXtract (MAX).

Though the TAF data represent a significant improvement in quality and usability over MAX, they remain highly complex, with varying data quality, eligibility categories, and data elements across states. As researchers begin to work with these data, there is an important opportunity to share learnings and approaches to avoid duplicative efforts and to distill key methodological standards. This effort will help ensure Medicaid research using the TAF is high quality, relevant, and impactful. With support from the Commonwealth Fund and the Robert Wood Johnson Foundation, AcademyHealth's Evidence-Informed State Health Policy Institute (ESHPI) established the Medicaid Data Learning Network (MDLN) to provide the opportunity to foster peer-shared learning among TAF users. Through a learning series curriculum, the MDLN provides a forum for TAF researchers to share what they have learned using the dataset and to develop consensus on best practices. These insights can then be disseminated to CMS, state Medicaid agencies, and the broader health services research community. The MDLN's ultimate goal is to improve the quality of the TAF data over time, expand opportunities for health services researchers to use Medicaid claims data, and increase the number of researchers engaged in Medicaid-focused research.

MDLN Design

At the core of MDLN's structure rests the project team, including AcademyHealth's ESPHI and three TAF research faculty, Drs. Sarah Gordon, John McConnell, and William Schpero. MLDN's Advisory Group is composed of stakeholders and experts with contextual experience, including members of AcademyHealth's Medicaid Outcomes Distributed Research Network (MODRN). Collectively, the MDLN project team designed a learning session curriculum. During Year One of the MDLN, this project team facilitated expert presentations and collaboration among TAF experts and member research teams by hosting an in-person meeting at AcademyHealth's Annual Research Meeting (ARM) and eight virtual learning sessions. These sessions, focused on priority topics identified by the MDLN research teams, created space for researchers working with TAF data to share their progress, questions, and solutions in a collaborative environment.

The learning session topics in Year One included:

- Approaches to Standard Measures of Utilization
- Linking MAX and TAF Data
- Navigating Race and Ethnicity Data in TAF
- Methods for Maternal and Reproductive Health Research Using TAF
- Measuring Managed Care Utilization Using Encounter Data in TAF
- Spending in Fee-For-Service and Managed Care Organization Delivery Systems
- Identifying Medications for Opioid Use Disorder in TAF

These learning sessions have proven to be a valuable resource in furthering understanding and enhancing the usability of TAF data among the research community. Furthermore, lessons from the MDLN sessions are relevant for the broader policymaking community at both the federal and state levels. Key takeaways for each group are explained below.

Learning Sessions

Approaches to Standard Measures of Utilization

Motivation

Health care utilization measures are the bread and butter of many claims-based research questions. Researchers may benefit from using standardized measures, which can aid in the reproducibility of studies and ensure that differences across studies or states do not arise due to different coding definitions.

Many utilization measures – such as primary care visits, emergency department (ED) visits, and inpatient stays – have common definitions, but still lack standardization across studies. Other measures, such as those produced by HEDIS, are proprietary, thus limiting their accessibility. Developing new utilization measures is even more difficult as these lack published definitions and accepted benchmarks to use for comparison and quality checks.

Considerations and Solutions

The MDLN covered several approaches to developing measures. One MDLN member team described their approach for identifying primary care utilization. They first pulled provider specialty from the Medicare Data on Provider Practice and Specialty (MD-PPAS) file, which reflects provider specialty in the Medicare Provider Enrollment, Chain, and Ownership System (PECOS). As a billing system, this file is checked often, and providers are incentivized to update their information every few years. (Notably, MD-PPAS will only be accurate for providers who are enrolled in Medicare. MD-PPAS might not be useful for those focusing on pediatric populations, for example). If MD-PPAS was missing, the team used

the TAF taxonomy code, if available. They used the code present on the plurality of other services (OT) claims. Finally, if the provider specialty was unable to be identified using the two methods listed above, the researchers used the primary specialty in the National Plan and Provider Enumeration System (NPPES) file. After identifying primary care providers (PCPs), the team used a broad definition of PCP utilization (any service provided by a PCP) and two narrower definitions: one defined by specific services provided by a PCP, and one slightly broader, defined by the same specific services but including other providers, such as OBGYNs.

A second MDLN member team compared three approaches for defining ED visits: the definition proposed by Research Assistance Data Center (ResDAC), a definition proposed in a 2008 paper by Handel and colleagues, and a HEDIS measure.[1] Among these three measures, the Handel and HEDIS measures were comparable, but ResDAC's definition produced notably lower estimates of ED visits.

A third team shared its work identifying services provided in rural health clinics (RHCs), a utilization measure that lacks an established definition and benchmark. The team identified several challenges, including grappling with states that may use different practices to code RHC services, missing procedure codes, and difficulties with reaching states' Medicaid offices for clarification, as the data in the TAF file differed from the state's own raw Medicaid and CHIP data.

The Data Quality (DQ) Atlas, developed by Mathematica, was highlighted as a critical starting point to understand the quality of certain utilization measures and identify high-level state exclusions, such as TAF not accurately reflecting a state's Medicaid enrollment. However, as a broad measure of quality, DQ Atlas may over- or understate the quality of a specific measure for a specific population and often requires further investigation based on a team's research question and population of interest. There are always exceptions; researchers should consider how their measures of interest might present new challenges. For example, prenatal care visits can be difficult to identify, as they are sometimes included in bundled payments, which can obscure prenatal visits for routine services.

As a group of researchers working with TAF, there is an opportunity to combine knowledge and develop standard definitions of key utilization measures. Expanding consensus on these definitions offers a variety of benefits, including enhanced reproducibility and credibility, common benchmarks that could be used to set quality standards, and a shared understanding of approaches to common challenges.

1 Handel, Daniel A., et al. "How Much Does Emergency Department Use Affect the Cost of Medicaid Programs." Annals of Emergency Medicine, Vol. 51, No. 5 (2008): 614-621.e1, doi: 10.1016/j.annemergmed.2007.09.002

Linking MAX and TAF Data

Motivation

When Medicaid launched in 1965, there was minimal data collection, and efforts to report individual-level data at the state or federal levels did not occur until the 1980s. In response to greater reporting needs, MSIS and MAX files were developed throughout the 1990s and early 2000s. Following the enactment of the Affordable Care Act (ACA), CMS transitioned from MSIS to T-MSIS, making TAF, rather than MAX, files available for research purposes in an effort to improve the quality, usability, and timeliness of national Medicaid claims data. TAF data are designed to represent an improvement over known quality issues in the MAX data. The transition from MAX to TAF occurred at different times in different states. In 2012, all states only used MAX, but by 2016, all states had transitioned to TAF.

Many data users are interested in using Medicaid claims to conduct longitudinal analyses that span the transition from MAX to TAF (2012-2016). The MAX/TAF transition spans significant program changes, including the 2014 Medicaid expansion, rapid growth in use of managed care, and expansion of behavioral health benefits. Depending on how many years of data researchers seek to use, availability varies by state and involves up to three different generations of data files.

Differences between MAX and TAF data create challenges in combining these generations of Medicaid data. In general, TAF includes more data than MAX. TAF has more than 1,400 data elements, compared to fewer than 400 in MAX. Data quality and oversight procedures also differ between the two data sources. In the MAX era, data were submitted quarterly, while TAF requires states to submit data monthly. Further, data quality procedures are automated across all states in TAF, which improves timeliness of review, whereas MAX data quality reviews occurred manually at the individual state level. CMS is currently working with states to fix data at the point of submission as opposed to after submission. While this process is slower, CMS is optimistic that the shift from MAX to TAF data will lead to a large shift from highly variable data quality to more consistent data quality with greater public transparency. Mathematica performs a variety of quality checks and assessments to provide a state-specific list of data quality issues to end-data users in the DQ Atlas.

Considerations and Solutions

In theory, individual-level beneficiary identifiers (IDs) should be consistent across MAX and TAF for the majority of records. Whenever states renumerate identifiers, they must provide CMS with a crosswalk to link the new IDs to the older-issue IDs. However, there are subsets of beneficiaries for whom linkages across MAX and

TAF are challenging due to state-specific changes to how MSIS IDs were assigned. In addition, there may be specific populations where linkage issues are more common for programmatic reasons, such as for newborns or enrollees with restricted benefits. Researchers with experience in longitudinal studies of MAX and TAF have experienced technical issues in linking data and saw attrition or loss of beneficiaries in the transition. Some changes in data elements may also present difficulties in linking across MAX and TAF because data elements in one file type do not have a perfect analog in the other file type, and vice versa. In addition, some variables may be present in both datasets but may be defined differently.

The usability of combined MAX and TAF data depends on users exploring, understanding, and sharing the quality of the linkages between the two files. CMS emphasizes the importance of communicating uncertainty as clearly and transparently as possible in any longitudinal study. Researchers should show each step of their analytic process, including reporting how many beneficiaries did not link across MAX and TAF, examining trends in study measures across the data file transition, and sharing when the file type transition occurred in each state used in the analysis. Data users should also pay close attention to the user guides for each file type. To grapple with policy implications based on longitudinal analyses, researchers must realistically discuss the interpretation of findings in the face of uncertainty, including what policymakers can do with the results and how actionable they might be.

Navigating Race and Ethnicity Data in TAF

Motivation

Medicaid programs do not require applicants or enrollees to self-report their race or ethnicity. Thus, race and ethnicity data in TAF are missing for a substantial proportion of enrollees, creating challenges to research assessing race- or ethnicity-based health disparities. The proportion of beneficiaries with missing race and ethnicity differs by state and by racial-ethnic group.

The DQ Atlas is a helpful resource for assessing the completeness of race and ethnicity data by state. Mathematica performs data quality checks, such as calculating the percent missing or invalid, assessing consistency between demographics observed in TAF compared to the American Community Survey (ACS), and computing indices of dissimilarity to examine consistency in states' data elements across months. These quality checks contribute to DQ Atlas' concern level for a particular state. States with a missingness value of 10% or less are reported as "low concern," 10-20% as "medium concern," and above 20% as "high concern." In 2017, 31 states were ranked as low or medium concern, 19 states and the District of Columbia ranked as high concern or unusable, and the territories of Puerto Rico and the Virgin Islands unclassified. The status in 2020 was similar, with 31 states and the territory of Puerto Rico ranked as low or medium concern, 20 states and the District of Columbia ranked as high concern or unusable, and the Virgin Islands remained unclassified.

Considerations and Solutions

Until the race and ethnicity data improve, researchers can consider several strategies for maximizing the usability of the data in TAF. First, they should conduct their own data quality investigations, particularly if they are focusing on a specific population. For example, a researcher's population of interest (e.g., children or pregnant women) may have a low level of missingness for race and ethnicity data in a given state, even if the DQ Atlas identifies the state as high concern. Researchers can compare rates observed in TAF to other data sources, including regional or county-level data, Medicare data, or project-related datasets such as the Centers for Disease Control and Prevention's (CDC) Natality Data to benchmark the racial and ethnic demographics observed in TAF.

Second, researchers can restrict their analyses to include only states that have adequate reporting of race or ethnicity data fields. However, states may have relatively low missingness for one racial group, but high missingness for a different racial group. To conduct cross-group comparisons, data users must limit their analyses to states that meet minimum thresholds for data quality for all racial-ethnic groups of interest.

Third, researchers may leverage panel or longitudinal data structures to impute a missing value for a specific beneficiary ID. Researchers experienced with TAF have indicated that this strategy has proven effective, citing that it may reduce observed missingness from 14% to 6%. In cases with conflicting race and ethnicity information across years, data users can employ "tiebreaker" strategies such as imputing the more common race and ethnicity or using data from the state that has better quality race and ethnicity data.

Finally, Mathematica is currently working on using Bayesian Improved Surname Geocoding (BISG) to impute race and ethnicity. BISG generates a predicted probability for each of six racial and ethnic categories for each person. Mathematica uses the Census Bureau's 2010 decennial Census and a Census-tract tabulation from the 5-year ACS to impute the BISG and has tested the method on TAF using numerous enhancements, such as limiting ACS data to the Medicaid population, regression analysis, adding first names, including an American Indian and Alaska Native (AI/AN) certification indicator, and enhancing calibration. CMS found that using first names and AI/AN certification offered the greatest improvements and modified the BISG to the Bayesian Improved First Name Surname Geocoding (BIFSG) with AI/AN certification and calibration. Mathematica anticipates that they will release imputed BIFSG estimates in the future to serve as a companion file for the TAF Research Identifiable Files (RIF). Mathematica advises that these

estimates be used for population-level analyses, rather than to make inferences about specific individuals. These data have been calculated for 2016-2019 data, and Mathematica is beginning to geocode 2020 data to apply the imputation.

Methods for Maternal and Reproductive Health Research Using TAF

Motivation

There are unique considerations when conducting maternal and reproductive health research using TAF. One such consideration is the approach for identifying births and dates of delivery in the Medicaid claims, both necessary first steps for research projects that seek to estimate maternal health or birth outcomes.

Considerations and Solutions

CMS has published two algorithms for identifying births in TAF on its website; other approaches have been used in prior studies that used MAX claims or state-specific Medicaid claims. Researchers have compared different approaches to identify live births in TAF, building on the CMS approaches. No research has evaluated approaches for identifying births in TAF and prior methods have relied on variables that are either unreported or have been restructured in TAF. One project used the 2018 TAF data to compare five approaches for identifying live birth counts and assessed their performance compared to the gold standard of birth record data from the CDC. They found that using a more restrictive code list excluding codes related to delivery services for infants achieved the best match of birth counts relative to CDC birth record data.

This work demonstrated that including claims from both the Inpatient (IP) and Other Services (OT) files and excluding codes unrelated to the delivery episode and those specific to services rendered to infants produced the best state-year match to natality data from the CDC. The comparison of different approaches to identify live births using the TAF provides a roadmap for researchers interested in using these data to answer questions related to maternal health. The approach used by researchers to identify live births may depend on the states of interest, as the performance of each approach varies across states.

It is common practice in some states for Medicaid programs to allow newborns to share a Medicaid identification number with their mother during their first year of life. This practice makes it difficult to separate billed services rendered to the mother versus the infant. One strategy to overcome this issue is identifying services or diagnoses that would only be relevant to a mother or a newborn (i.e., jaundice) and either restricting to only those measures, or using the presence of services rendered to both a mother and an infant as a basis for excluding that beneficiary ID due to the presence of sharing. Alternatively, data users can exclude states in which there is known to be a high proportion of ID sharing based on a previously published CMS report.

Given that the TAF are new and relatively unvalidated, it is important to benchmark maternal and reproductive quality measures to external data sources when possible. When no external benchmark exists, data users should plot mean outcomes among all states to identify outliers and compare individual state measures to the national median.

Measuring Managed Care Utilization Using Encounter Data in TAF

Motivation

As of July 1, 2020, 84% of Medicaid beneficiaries were in some managed care program, and 72% were in comprehensive managed care (CMC) programs. States are slowly continuing to expand eligibility for managed care, increasingly including beneficiaries with complex needs.

The growth in managed care has implications for measuring utilization. Under the traditional fee-for-service model, health care providers are paid for each service they provide. Under the managed care model, however, providers may receive different types of payments, including capitation, which provides a per-enrollee payment that does not increase or decrease based on the number of visits. Visits for enrollees in Medicaid managed care are reflected in TAF as "encounter" records.

The reporting of these encounter claims by managed care organizations (MCOs) did not become mandatory until after the passage of the ACA in 2010. Although most plans and states comply, the approaches have not been completely standardized, allowing for differences across states and plans. Thus, researchers interested in working with TAF to assess utilization may need to adjust their approaches to ensure reliable outcomes.

Considerations and Solutions:

Researchers should start by assessing the quality of managed care data by reviewing the DQ Atlas, which summarizes the quality of managed care utilization data as low concern, medium concern, high concern, or unusable. If a state is flagged as problematic in the DQ Atlas, it is best to investigate the data independently to determine its overall quality. The DQ Atlas assesses enrollment counts for comprehensive managed care, primary care case management programs, and behavioral health organizations. TAF includes those data, as well as enrollment in capitated dental plans, transportation plans, the Program of All-Inclusive Care for the Elderly (PACE), and other non-comprehensive prepaid health plans. In addition to reviewing DQ Atlas, researchers may want to supplement their inclusion and exclusion criteria based on an empirical review of their own data.

Researchers interested in utilization within specific managed care plans may need additional work to place enrollees within an MCO. TAF data include an annual managed care plan file (APL) which may be helpful in these endeavors. However, the APL files have

not yet been used extensively. With respect to managed care plan IDs, researchers may need to read through the coding guidance CMS published for state reporting of managed care plan IDs in enrollment records. The guidance provides key examples of how state IDs for their plans can make it challenging to identify specific plans. Researchers should use the managed care plan type code in conjunction with the plan ID when identifying individual plans. This guidance assumes that most studies will want to focus on specific plan types. The managed care plan data have not been used extensively, so researchers should proceed with caution.

Spending in Fee-For-Service and Managed Care Organization Delivery Systems

Motivation

In 2019, spending in the Medicaid program totaled approximately $613.50 billion, along with $19.9 billion in CHIP. Managed care capitated payments accounted for more than half of total expenditures; as such, accurately measuring spending in the TAF is critical for researchers. These capitated payments that states make to insurance plans are the dominant part of expenditures in Medicaid ($310 billion), followed by long-term care, physicians, labs and other services, inpatient services, administrative, and prescription drugs. When divided into 5 major eligibility categories (children, adults, Medicaid expansion adults, adults with disabilities, and the aged), per capita expenditures vary widely. The aged and adults with disabilities had the largest expenditures on a per capita basis, while children and other adults make up the majority of Medicaid enrollees. Though using payment data to measure spending in TAF can be difficult, some solutions are available.

Considerations and Solutions

The CMS-64 reports can provide a useful benchmark for spending estimates derived from the TAF. CMS-64 data are pulled from the Medicaid Budget and Expenditure System/State CHIP Budget and Expenditure System (MBES/CBES), which is the financial reporting system used to provide federal matching funds to states, who report into the system on a quarterly basis. However, this reporting is based on what states paid in the previous quarter and may not correspond to the actual services rendered. Though the CMS-64 data are aggregate, it is possible, for example, to use the data to estimate inpatient expenditures by quarter by state, but the CMS-64 data do not map to the granular level of claims service. TAF claims represent the opposite end of the spectrum of granularity, as these claims capture data at the service level. Expenditure data include fee-for-service (FFS), capitated payments, encounter claims, supplemental claims, and service tracking records. Unfortunately, both supplemental claims and service tracking records are likely highly underreported. When benchmarked expenditures from TAF data are compared to CMS-64, from 2018-2021, TAF expenditures are 15% less than CMS-64. Notably, the paid amount on encounter claims reflecting payments from managed care plans to providers are redacted in the TAF.

TAF data can most easily be mapped to CMS-64 data using the Title 19 code (XIX_SRVC_CTGRY_CD). More recent versions of the TAF include a federally assigned service category (FASC), which facilitates apples-to-apples comparisons of utilization and spending for defined populations across states.

The TAF data include five broad types of payments:

- FFS Expenditures – These are tied to a beneficiary for a specific service and are identified as claim type 1 for Medicaid and claim type A for S-CHIP in T-MSIS. There is a DQ Atlas topic for data quality of FFS expenditures and different categories.
- Managed Care Capitation Payments – These and other monthly payment amounts are tied to a beneficiary for all medical services. They are identified as claim type 2 for Medicaid and claim type B for S-CHIP in T-MSIS. There is a DQ Atlas topic for data quality of capitated payments.
- Managed Care Companies to Providers – These are encounter payments and are identified as claim type 3 for Medicaid and claim type C for S-CHIP in T-MSIS. They are redacted and proprietary.
- Financial Transactions – These are considered "supplemental payments," although this term can be confusing because it is often associated with upper payment limits (UPL). However, in TAF, they refer to additional costs for a beneficiary, such as when a Federally Qualified Health Center (FQHC) receives a "supplemental payment" above the standard amount received for a service for a beneficiary. These are identified as claim types 4 and 5 for Medicaid and claim types D and E for S-CHIP in T-MSIS.
- Other – Other payments are identified as claim types U, V, W, X, Y, and Z for S-CHIP in T-MSIS. States are not using these claim types as frequently as they once were. U is "other FFS claims," V is "other capitation claims," W is "other managed care encounters," X is "non-Medicaid or CHIP service tracking claims," Y is "other supplemental payment," and Z is "denied claims."

Identifying Medications for Opioid Use Disorder in TAF

Motivation

Medicaid is the largest payer of substance use disorder services in the United States and therefore many researchers have an interest in using TAF data to study substance use disorders and treatment. However, codes and methods for identifying cohorts, diagnoses, utilization, and outcomes have not been standardized. Furthermore, large differences in state systems and Medicaid coverage of substance use can make it difficult to compare or benchmark across states. The MDLN devoted a session to address some of the challenges specific to two measures relevant to substance use disorder: injection drug use and methadone use.

Considerations and Solutions

Identifying Injectable Drug Use. There is no established method for identifying people who inject drugs in claims data. In one study, researchers looked for any instance of any code in the following categories: opioid abuse/dependence; opioid poisoning; abuse/dependence or poisoning for other injectable drugs; prescription for buprenorphine; Healthcare Common Procedure Coding System (HCPCS) code consistent with methadone maintenance treatment, using a 12-month lookback period from the date of diagnosis.[2] The study reported sensitivity ranging from 65-90% and specificity of 47-79%.

This approach still has limitations, including its inability to identify changing drug use patterns, such as non-injection prescription opioid use and an inability to distinguish "current" or ongoing injection drug use from prior injection drug use. In addition, benchmarks are not yet well-established to assess quality.

Identifying Methadone Use. Researchers with experience identifying methadone use in TAF used codes H0020 and S0109 in conjunction with three variables that capture procedure codes: line_prcdr_cd (Line procedure code); ot_accmdtn_hcpcs_rate (Other Services Accommodation Rate); and prcdr_cd (Procedure Codes 1-6).

In their review of data, researchers noted that sixteen states and a territory did not have any methadone claims (ID, ND, IL, WY, SD, UT, NE, KY, KS, AR, TN, SC, OK, LA, MS, PR). However, during the time these data were collected, not all states covered methadone, including AR, ID, KS, KY, LA, NE, ND, OK, PR, SC, TN, and WY. Researchers used a variety of sources to identify states that did, in fact, cover methadone, and to clarify why 14 of the 16 states had zero methadone claims. Research implications of data quality concerns depend on the research question and the type of missing data. Researchers also noted that two states (SD and MS) covered methadone but had no Opioid Treatment Programs (OTPs) that provided methadone.

Takeaways for the Research Community

The TAF data represent a significant improvement in quality and usability over previous data. While they remain highly complex, with varying data quality issues across states, researchers have begun to learn how to use these data. Indeed, studies using TAF data have now appeared as published manuscripts, and researchers in the MDLN have begun sharing new approaches and lessons learned from unpublished work.

2 Janjua, Naveed Zafar, et al. "Identifying Injection Drug Use and Estimating Population Size of People Who Inject Drugs Using Healthcare Administrative Datasets." The International Journal on Drug Policy, Vol. 55 (2018): 31-39. doi:10.1016/j.drugpo.2018.02.001

The next few years should bring a significant increase in TAF studies that will give policymakers greater insights into the Medicaid program. In our first year of the MDLN program, we offer these takeaways for the research community:

- **Studies of utilization**, including regional variation, are underway and appear reliable, particularly when relying on common utilization measures and validation via external benchmarks.
- **Future research** will likely feature longitudinal studies that use TAF and MAX data together but will require significant efforts to validate observed trends and address measurement uncertainty.
- **Studies that rely on race and ethnicity data** may be limited in the near future because race and ethnicity data are missing or unusable for many states. However, imputation approaches may lead to new files released by Mathematica or CMS that help bridge this gap.
- **While managed care utilization data appear** reliable, researchers have not yet begun to unpack the managed care enrollment files in ways that would allow them to compare managed care types or different managed care approaches.
- **In some instances, TAF data can be linked to other data sources to enrich analyses.** Agreements are being developed that would allow for datasets that link TAF data with the National Health Interview Survey (NHIS), National Health and Nutrition Examination Survey (NHANES), and the National Hospital Care Survey (NHCS). These linked data are currently available through the Research Data Center (RDC) network. Interested researchers must submit a research proposal to the RDC. CMS has recently revisited its policy regarding the release of identifiers for linkage with the National Death Index (NDI) data. Currently, CMS does not provide direct identifiers to researchers and cannot provide identifiers to NCHS on behalf of researchers. CMS and NCHS are working to develop a process to facilitate external researchers receiving NDI data from NCHS, but the timeline for implementation is unclear.

Takeaways for the Policy Community

Since the first release of TAF in 2019, the data have helped to catalyze policy-relevant research on the Medicaid program. Yet, the MDLN has identified several opportunities to improve the usability of certain data elements and the accessibility of the dataset. Specific opportunities are as follows:

- **Race and Ethnicity Data Quality**. Both the MDLN and a variety of other organizations have documented significant gaps in the quality of race and ethnicity data in the TAF, which limits the ability of researchers to study health equity in the Medicaid program. These data quality issues are multifactorial in origin, stem-

ming in part from (i) missing race and ethnicity data on enrollment forms (beneficiaries are not required by law to provide this information at the time of enrollment); and (ii) underlying data quality issues involved in state- and federal-level processing of T-MSIS. To address these gaps, we suggest CMS work with states to develop best practices to encourage beneficiaries to provide race and ethnicity data at the time of enrollment, as well as provide additional resources for states to invest in this important data collection infrastructure.

- **Spending on Encounter Claims**. Reimbursement from MCOs to providers is redacted in the TAF data, although evidence suggests that the underlying quality of the reimbursement data is also low. Given that more than 80% of Medicaid beneficiaries are now enrolled in some form of managed care, access to high-quality spending data is critical for understanding the value delivered by MCOs and how providers respond to financial incentives in the Medicaid program. One option is for CMS to mandate the collection of high-fidelity payment data from all MCOs and promulgate new regulations that allow for the full release of MCO payment data in the TAF. This step is critical for ensuring transparency and accountability in Medicaid.
- **Data Accessibility**. TAF data files are large, complicated to use, and expensive. Buying a full year of physical TAF data costs over $65,000; access to the data files via the CMS Virtual Research Data Center (VRDC) is more than $30,000 annually. As a result, research using the TAF data is cost-prohibitive for many organizations. CMS can democratize access to the TAF data and diversify the set of organizations engaged in policy-relevant Medicaid research by significantly lowering TAF access fees or offering them on a sliding-fee scale based on organization type and resourcing. In addition, while the timeliness of TAF releases has greatly improved relative to the predecessor MAX files, the data are still only available at a significant lag. We recommend CMS work with states and relevant contractors to ensure that the first release of TAF for a given year is available no later than one year later.
- **Data Transparenc**y. The TAF data are a work in progress as CMS, its contractors, and states attempt to improve data quality and reliability. Although significant documentation is available for TAF users (most notably through the DQ Atlas), there is limited transparency and documentation on the timing of data releases and details on new data elements as they are added over time. CMS could improve documentation on TAF by posting a clear timeline of data releases and detailed release notes that delineate changes to and improvements in the data files over time.

Conclusion

The TAF represents a powerful opportunity to conduct timely, policy-relevant research on the Medicaid program. However, findings from the first year of the MDLN highlight important complexities and data quality challenges that limit the effectiveness of the dataset for certain research questions. Going forward, the MDLN seeks to (i) engage in bidirectional feedback with state and federal policymakers to improve data quality, timeliness, and accessibility; and (ii) collaborate with the TAF research community, leveraging insights from the first year of the learning network, to develop standardized approaches for TAF-derived research to promote transparency and replicability.