MEDICAID CLAIMS DATA: USING T-MSIS FOR STATE HEALTH OUTCOMES RESEARCH

AUTHORS

Quin Mudry Nelson, MPH
Health Equity Fellow &
Research Assistant

Claire McGlave, MPH
PhD Student

What is T-MSIS?

Curated by the Centers for Medicare and Medicaid Services (CMS), T-MSIS, or the Transformed Medicaid Statistical Information System, is a comprehensive database of claims made to state Medicaid and CHIP programs by their beneficiaries. This includes data from all 50 states and the District of Columbia.

T-MSIS data include claims from:

- inpatient (in-hospital) services
- outpatient services
- long-term care, and
- prescription drug and/or pharmacy services

T-MSIS claims are available starting in 2014. There is an 18-month lag between when claims are filed and when they are published in the T-MSIS dataset. CMS updates the data quarterly, publishing preliminary and final versions of the data.

T-MSIS File Structure

The T-MSIS analytic files are divided into seven parts: inpatient claims, long-term care claims, pharmacy claims, claims for "other" (i.e., outpatient) services, a patient demographic file, a patient eligibility file, and an annual provider file (see Table 1).

Table 1: T-MSIS Data Files by Type

File	Explanation
Inpatient Files	Records on encounters that occur in the hospital
Long-Term Care Files	Records of visits to nursing facilities, mental health facilities, and psychiatry units. Includes intermediate care for individuals with intellectual disabilities
Pharmacy Claims	Records of drugs prescribed and filled at a pharmacy, and other services provided by pharmacists (e.g., vaccinations)
Other Files	Records of claims, encounters, and payments that took place in the outpatient setting (physician services, dental services, labs, imaging, home health services, durable medical equipment, and others)
Patient Demographic File	Information on patients (e.g. sex, age, geographic location, etc.)
Patient Eligibility File	Enrollment information about patients, specifically which Medicaid plan they were enrolled in and when
Annual Provider File	Information on providers (e.g. sex, age, licensure, where they practice)

Source: CMS T-MSIS Files (2024)

Each of these files can be linked to yield information on the services and drugs provided at each visit, demographic information about the patient, patient plan details, and background on the provider and where they practice.

Why is T-MSIS Data Useful?

T-MSIS data is useful because it is a comprehensive record of services rendered to Medicaid beneficiaries across the country. Specifically, T-MSIS files include data from every Medicaid program in the United States – programs from all 50 states and Washington D.C. report claims data that are harmonized into a national database. Additionally, T-MSIS includes data from U.S. territories Puerto Rico and the U.S. Virgin Islands, with plans to integrate Guam's data in the coming years.

T-MSIS data are collected, cleaned, and harmonized by CMS. This gives researchers access to a single database that includes information from many Medicaid programs, making national-level analyses possible. This results in a large dataset of decent quality for assessing and comparing utilization, spending, and health outcomes across different state Medicaid programs. This dataset represents significant improvements in quality and comprehensiveness of Medicaid claims data. However, it is important to note that this data has varying levels of completeness (see "Quality Concerns with T-MSIS Data").

How Researchers Have Used T-MSIS Data

Thanks to the wide variety of cleaned and harmonized data along breakdowns by state, specific plan details, provider details, and breakdowns by demographic categories, T-MSIS claims data can be used to evaluate an almost unlimited range of questions.

Here are just a few examples of how researchers have used T-MSIS data in their work:

- McConnell and co-authors examined rates of emergency department visits for psychiatric conditions among Medicaid beneficiaries across the U.S. Using the T-MSIS data allowed them to differentiate between patient and place-specific factors to disentangle factors that lead to higher emergency department use.[1]
- Meinhofer et al. used the T-MSIS data to characterize the use of opioids and stimulants among pregnant Medicaid beneficiaries, finding that the use of opioids has decreased, but the use of stimulants has increased. This provides crucial information to Medicaid programs about how to provide the best care for their patients.[2]
- Forrest and co-authors compared claims for pediatric care in T-MSIS to claims for commercially insured children. They found that children insured by Medicaid were less likely to see a specialist in the outpatient setting compared to their commercially insured counterparts. These findings expose important differences in quality of and access to care for Medicaid beneficiaries.[3]

Quality Concerns with T-MSIS Data

While the T-MSIS data represent dramatic improvements in the quality, comprehensiveness, and harmonization of Medicaid administrative data, the lack of completeness that results from harmonizing data from over 50 distinct insurance programs poses significant challenges to data quality.

Chief among these is the quality of reporting data on patient characteristics. While many Medicaid programs have improved the quality of their data collection over time, individual states' programs vary in their ability to report several key variables, namely race and ethnicity.

In response to concerns over data quality issues, CMS has created the T-MSIS Data Quality Atlas for users of this data to view the completeness of almost every variable in the data set by year. This enables researchers to understand which variables are usable before requesting data access.

---

[1] McConnell, J., Watson, K., Choo, E., and Zhu, J.M. "Geographical Variations in Emergency Department Visits for Mental Health Conditions for Medicaid Beneficiaries" *Health Affairs (2023).*

[2] Meinhofer, A., Martinez, M.L., and Palmsten, K. "Patterns in Prescription Opioids, Benzodiazepines, and Stimulants Filled by Pregnant Medicaid Beneficiaries" *JAMA HealthNetwork (2022).*

[3] Forrest, C.B., et. al, "Pediatric Medical Subspecialist Use in Outpatient Settings" *JAMA Network Open* (2024).

How Can T-MSIS Be Improved to Promote Health Equity?

T-MSIS is a great data resource for comparing outcomes between state Medicaid programs across basic demographics such as age, gender, and broad racial categories. The data also give researchers the tools to understand the care Medicaid beneficiaries receive, and how that has changed over time.

The data quality concerns we mentioned earlier limit the micro-level use of this data for policy goals like promoting health equity. However, the T-MSIS data is undergoing significant improvements that will lead to better representations of various demographic groups, smaller communities, and marginalized groups.

One way that CMS is working to improve this dataset is via the Data Quality Atlas mentioned above. The Atlas gives data users information about which data elements are incomplete, but also signals reporting weak points to CMS. Understanding where gaps exist gives CMS the tools to direct resources to states that need help improving their reporting.

One of the most acknowledged ways in which all health data sources, including T-MSIS, can be improved to promote health equity, is to disaggregate data categories. More specific categorizations and breakdowns can help researchers and practitioners better understand the variability within the more general and overarching racial and ethnic groups.

With such a large, population-level dataset like the T-MSIS, implementing more specific categories and breakdowns for identity will enable researchers to understand the diverse experiences of people under one racial identity.

An example of this idea can be seen when looking at disaggregated data of those who identify as a part of the Asian American Pacific Islander (AAPI) racial category. There is a significant difference in socioeconomic status, health status, income, employment status, and many other measures among different ethnicities within the AAPI racial category, as highlighted in this 2016 report from the White House Initiative on Asian American and Pacific Islanders Interagency Working Group Data and Research Committee.

These significant differences that exist within this specific group of people would not have been visible without the implementation of additional categories for identity. Thus, increasing the options for self-identification could reveal differences and gaps within communities and groups.

Addressing the data quality concerns (in addition to improving upon the categories in which data is disaggregated) can expand the usage of T-MSIS as a state and federal data source for health outcomes research. These changes are already being implemented at the federal level through the Office of Management and Budget's Statistical Policy Directive No. 15, which has been revised to combine all race and ethnicity questions into one comprehensive question where respondents can choose as many options as they identify with, as well as an expansion and revision to the language of certain racial and ethnic groups.

Learn more about how these recent federal level changes may affect state level data collection and promote equity from this webinar hosted by State Health & Value Strategies (SHVS) with SHADAC and Health Equity Solutions.

How Do You Access T-MSIS Data for Research?

The Application Process & Estimated Timeline

Now that you know what T-MSIS data is, its limitations, and how it is used, you can decide whether this dataset will be useful and applicable for your own research. If you do want to use this dataset, you'll need to apply for access using the following process:

The Research Data Assistance Center (or ResDAC) at the University of Minnesota administers the application process for T-MSIS data for CMS. The application process consists of several steps and can take anywhere six to eight months from initial submission to ResDAC to accessing data on the Virtual Research Data Center (VRDC) (see Figure 1), depending on the complexity of the application. The following outlines the steps for applying for access to T-MSIS data for research:

1. Prepare to apply

To apply, investigators must seek a human subjects exemption from their institution's IRB (Institutional Review Board).

2. Reach out to the Research Data Assistance Center (ResDAC)

Use ResDAC's request forms generator to make a list of the required forms. These (generally) consist of:

- a short protocol for the study
- a list of study team members, and
- a specification of the study sample and variables needed

3. Seek additional HIPAA waiver from the IRB

Before an application for T-MSIS data is submitted to CMS, investigators must also certify a data-only (or specimen-only) protocol with their institution's IRB (this is sometimes called a "HIPAA waiver").

4. ResDAC reviews application materials

ResDAC will provide feedback on study protocols and data request forms, which investigators will likely need to revise based on ResDAC feedback in order to meet CMS requirements.
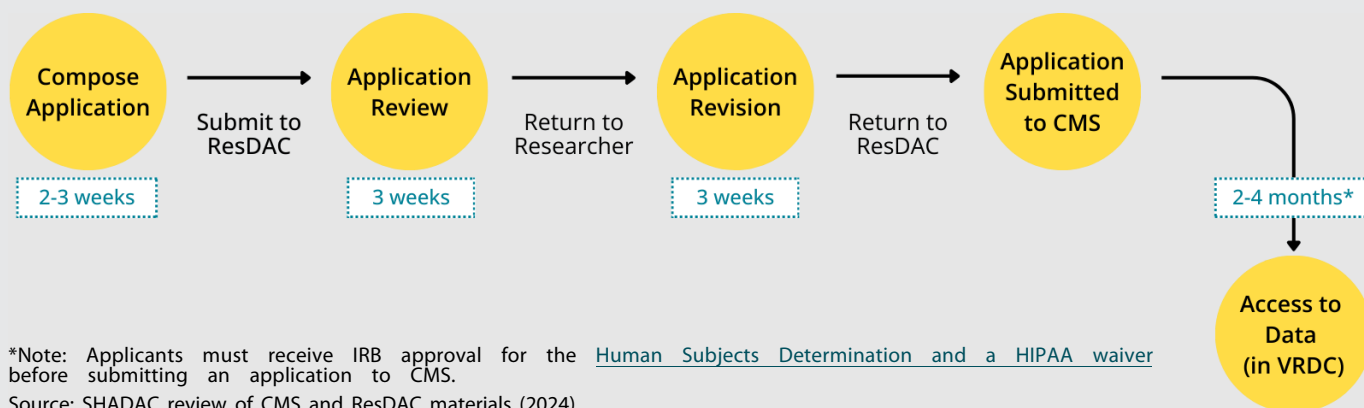
5. ResDAC submits final application to CMS

ResDAC will review investigator edits and submit the data application to CMS on behalf of the study team. Note that the approval process can take several months to complete in total – so if you plan to use T-MSIS data for a study in the near future, plan for a six-to-eight month application process (including communication with ResDAC, waiting time, and CMS processing time).

Additional considerations

T-MSIS files can be cost-intensive (see ResDAC's cost estimates for further information). This cost, however, covers the attainment of a 100% sample of claims data. Once a data cleaning infrastructure is established, researchers can file Data Use Agreements (DUAs) with CMS to reuse the same data for additional research questions and studies.

Figure 1: T-MSIS Data Application Timeline

*Note: Applicants must receive IRB approval for the Human Subjects Determination and a HIPAA waiver before submitting an application to CMS.
Source: SHADAC review of CMS and ResDAC materials (2024)

Conclusion

T-MSIS is a relatively novel and extremely rich dataset that gives researchers the tools to understand Medicaid beneficiaries and the providers who care for them. It enables the study of practice patterns and the comparison of Medicaid programs over time. CMS' ongoing efforts to improve data quality and comprehensiveness are publicly available and allow researchers to explore data limitations before applying for data access.

Despite room for improvement in some areas of demographic measurement (e.g., race and ethnicity), T-MSIS is a valuable resource for researchers to understand the administration and outcomes of care for the Medicaid population.