The Quality of Race and Ethnicity Data Among Medicaid Beneficiaries Receiving Long-Term Services and Supports

Laurin Bixby, Elad Daniels, Joe Caldwell

CLPC
Community Living Policy Center

Community Living Equity Center
Centro para la Equidad en Vida Comunitaria

Brandeis THE HELLER SCHOOL FOR SOCIAL POLICY AND MANAGEMENT
Lurie Institute for Disability Policy

Background

The Transformed Medicaid Statistical Information System (T-MSIS) provides state-reported Medicaid claims data for the population of Medicaid beneficiaries across U.S. states. T-MSIS data is critical for developing evidence-based practices to improve Medicaid services, including Long-Term Services and Supports (LTSS). Researchers can use T-MSIS data to identify disparities in access to and utilization of services across states and demographic characteristics, including age, disability, rural/urban, sex, and race-ethnicity. This information can be used to improve access, quality, and outcomes of Medicaid services.

Though T-MSIS is the most comprehensive data on Medicaid beneficiaries that is available, there are known concerns regarding the completeness and quality of T-MSIS data. The Data Quality (DQ) Atlas provides information on the quality and usability of different topics in the T-MSIS data, including beneficiary information such as race and ethnicity. The DQ Atlas assesses the quality of a state's race and ethnicity data in T-MSIS based on completeness (the percent of Medicaid beneficiaries in a state that are missing race and ethnicity data) and consistency with race and ethnicity data among Medicaid recipients for that state in the American Community Survey. Based on this assessment, the DQ Atlas designates states as having low concern, medium concern, high concern, or unusable data. While the quality of T-MSIS race and ethnicity data has improved over time, there are still many states with high concern and unusable data.

Given the race-ethnicity data quality concerns in T-MSIS, many researchers and the Centers for Medicare & Medicaid Services (CMS) use imputed race and ethnicity data to report racial and ethnic differences in service utilization. Imputation involves assigning someone's race and ethnicity based on information such as an enrollee's name and zip code. Though the practice of imputing race and ethnicity has been a long-established practice, many advocates advise against this practice as it raises major statistical and ethical issues and can create biased, inaccurate information. From an equity perspective, researchers need to be using self-reported race and ethnicity data rather than imputing race and ethnicity data.

Currently, the DQ Atlas does not include an assessment of LTSS data. However, CMS's LTSS expenditure and user reports often exclude certain states due to LTSS data quality concerns. For example, in 2021 these reports exclude Alabama due to concerns regarding the quality of LTSS data in T-MSIS. Though they are not excluded from CMS reports, many states have expressed additional concerns regarding the quality and accuracy of their LTSS data in T-MSIS. It is important to understand data quality concerns as we use and interpret T-MSIS data. Additionally, a more extensive assessment of LTSS data quality should be included in the DQ Atlas so that researchers can contextualize T-MSIS information on Medicaid LTSS and advocates can push for improved data collection.

Researchers using the T-MSIS data to look at race and ethnicity among the LTSS population must be aware of the quality and limitations of the T-MSIS data or else risk putting out biased and inaccurate information. In the present brief, we examine the completeness of race and ethnicity data among adult Medicaid LTSS users, including adults receiving Home and Community-Based Services (HCBS) and institutional services. While the DQ Atlas assesses the quality of race and ethnicity data among all Medicaid beneficiaries, we focus our analysis on the quality of race and ethnicity data among adults who receive Medicaid LTSS.

> *Researchers using the T-MSIS data to look at race and ethnicity among the LTSS population* ***must be aware of the quality and limitations of the T-MSIS data or else risk putting out biased and inaccurate information.***

Methods

Assessment of T-MSIS Race and Ethnicity Data Among LTSS Users

We evaluated the quality of race and ethnicity data among adults receiving LTSS in the 2021 T-MSIS based on the completeness of the data. Specifically, we examined the proportion of a state's adult Medicaid LTSS population that is missing race and ethnicity data. We calculated the percent of missing race and ethnicity data among adults receiving HCBS (n=5,024,784) and among adults receiving institutional LTSS (n=1,124,026). In the Appendix table, we also present the percent of missing race and ethnicity data among adult Medicaid beneficiaries who do not receive LTSS (n=53,483,044).

We determined who receives HCBS and institutional LTSS using CMS's methodology, which is based on the definition of HCBS in section 9817 of the American Rescue Plan Act of 2021. HCBS includes 1915(c) Waiver Programs, 1915(i) HCBS State Plan Option, 1915(j) Self-Directed Personal Assistance Services, 1915(k) Community First Choice, Program of All-Inclusive Care for the Elderly, Personal Care Services, Home Health Services, Rehabilitation Services, Case Management Services, and Private Duty Nursing Services. Institutional LTSS includes services provided in nursing facilities, mental health facilities, or Intermediate Care Facilities for Individuals with Intellectual Disabilities.

In our data quality assessment, we excluded adults who received both HCBS and institutional LTSS in the same year, which is 3.64% of adult Medicaid beneficiaries receiving LTSS. Additionally, we cannot assess the quality of race and ethnicity data for Alabama or Mississippi due to concerns identified by CMS regarding the quality of Alabama's LTSS data and missing age information in the T-MSIS demographic and eligibility file for Mississippi.

We followed the methodology used by CMS in the DQ Atlas to classify states as either low concern, medium concern, high concern, or unusable based on the percent of missing race and ethnicity data among LTSS users in each state. We designated a state as *low concern* if they were missing less than or equal to 10% of race and ethnicity data, *medium concern* if they were missing 10 to 20% of race and ethnicity data, *high concern* if they were missing 20 to 50% of race and ethnicity data, and *unusable* if they were missing 50% or more of their race and ethnicity data. The DQ Atlas additionally bases their data quality assessment on the accuracy of race and ethnicity data in T-MSIS in comparison with the racial and ethnic distribution of Medicaid recipients in the American Community Survey. We are unable to account for the accuracy of race and ethnicity in this way because the American Community Survey cannot identify a comparable comparison group of Medicaid LTSS recipients.

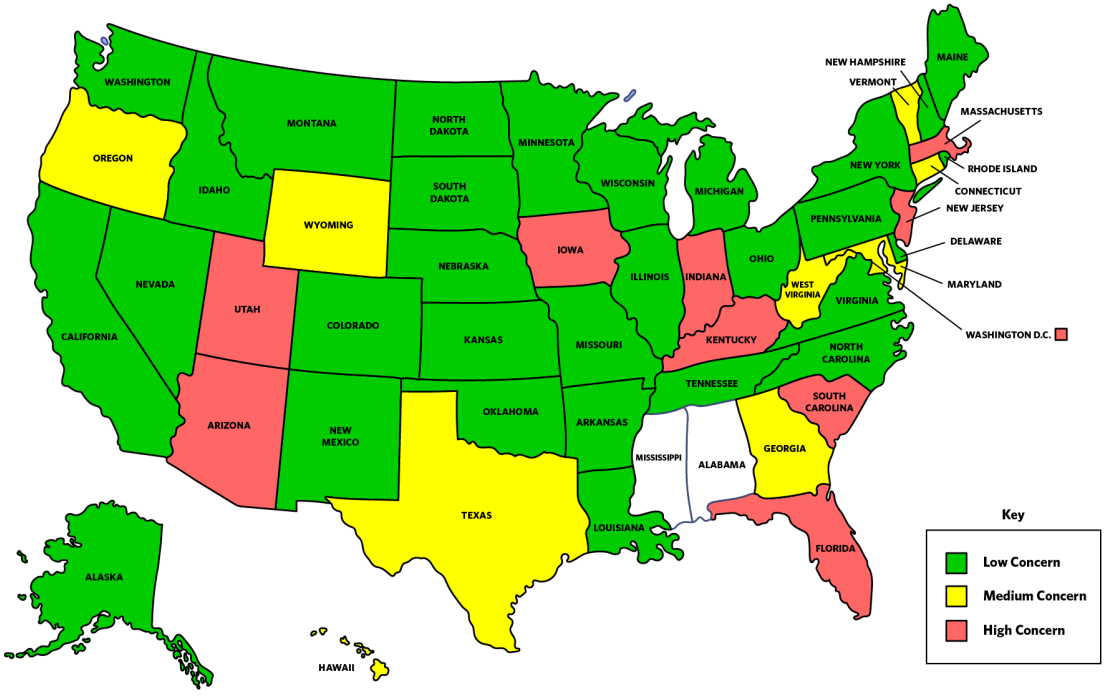
Findings

Quality of Race and Ethnicity Data among HCBS Users

The map below illustrates the results of our race and ethnicity data quality assessment among adults receiving Medicaid HCBS. Here we find that there are ten states deemed high concern, meaning that 20-50% of race and ethnicity data is missing. These states include Arizona, Florida Indiana, Iowa, Kentucky, Massachusetts, New Jersey, South Carolina, Utah, and Washington D.C. We recommend excluding these states in any analysis related to race and ethnicity among Medicaid HCBS users.

There are nine states designated as medium concern among adults receiving Medicaid HCBS. Approximately 10-20% of adults in Connecticut, Georgia, Hawaii, Maryland, Oregon, Texas, Vermont, West Virginia, and Wyoming are missing race and ethnicity data. The remaining states, excluding Mississippi and Alabama, are designated as low concern, meaning that 0-10% of adults receiving Medicaid HCBS are missing race and ethnicity data in these 30 states. Since our assessment is only able to speak to the completeness of the data, we advise caution even with data from low concern states because there may be concerns with the accuracy of data reporting.

Race and Ethnicity Data Quality Assessment among Adults Receiving Medicaid HCBS



The data quality assessment is based on the percent of missing race and ethnicity data in the 2021 T-MSIS Analytic File (Low concern: ≤10%; Medium concern: 10% - ≤20%; High Concern: 20% - ≤50%; Unusable: >50%).

Alabama is excluded due to concerns identified by CMS about the quality of Alabama's LTSS data.
Mississippi is excluded due to missing age information in the demographic and eligibility (DE) TAF.

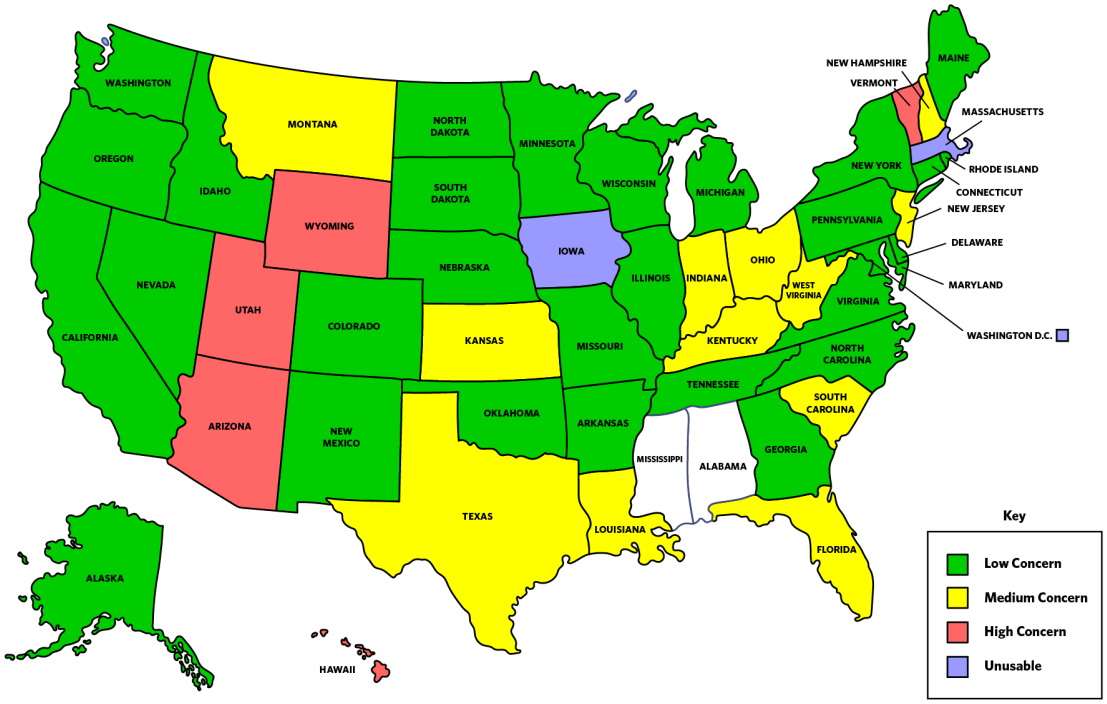
Quality of Race and Ethnicity Data among Institutional LTSS Users

The map below shows the race and ethnicity data quality assessment results among adults receiving Medicaid institutional LTSS. There are three states with over 50% of their institutionalized adult population missing race and ethnicity data. Specifically, Iowa, Massachusetts, and Washington D.C. have unusable data because they are missing over half of their race and ethnicity data among adults receiving institutionalized LTSS. There are five additional states with high concern data: Arizona, Hawaii, Utah, Vermont, and Wyoming. We advise against using data from states with unusable or high concern race and ethnicity data.

Twelve states have medium concern race and ethnicity data among adults receiving institutional LTSS, including Florida, Indiana, Kansas, Kentucky, Louisiana, Montana, New Hampshire, New Jersey, Ohio, South Carolina, Texas, and West Virginia. The remaining 29 states, excluding Alabama and Mississippi, have low concern race and ethnicity data among adults receiving institutional LTSS. As above, we still recommend being cautious when using

data from medium and low concern states as there may be inaccuracies in the race and ethnicity data.

Race and Ethnicity Data Quality Assessment among Adults Receiving Medicaid Institutional LTSS



The data quality assessment is based on the percent of missing race and ethnicity data in the 2021 T-MSIS Analytic File (Low concern: ≤10%; Medium concern: 10% - ≤20%; High Concern: 20% - ≤50%; Unusable: >50%).

Alabama is excluded due to concerns identified by CMS about the quality of Alabama's LTSS data.
Mississippi is excluded due to missing age information in the demographic and eligibility (DE) TAF.

Discussion

Why is the quality of race and ethnicity data in T-MSIS bad?

There are several reasons why the quality of race and ethnicity data in T-MSIS is inadequate. One possible reason is issues with data infrastructure. Moreover, it is possible that race and ethnicity data is being collected in medical records or other administrative files, but it is not being reported into the claims system and is thus missing in T-MSIS. It is also possible that health plans are collecting this data and reporting it to states but states may be overriding this.

Another reason for the substantial missing data in T-MSIS is that individuals may not be reporting their data, such as on the Medicaid application, due to distrust in how the

government will use this data. In addition to improving data infrastructure across states, it is critical to provide clear explanations of *why* race and ethnicity data is being collected and how it will be used. The race and ethnicity data in T-MSIS is also of poor quality due to the particularly inadequate collection of data on multiracial individuals. It appears that many states do not include multiracial as a race and ethnicity category in their data collection efforts. There needs to be more consistency across states in how they collect race and ethnicity data in the T-MSIS system.

> *In addition to improving data infrastructure across states, it is critical to provide clear explanations of **why** race and ethnicity data is being collected and how it will be used.*

Policy Implications

The quality of race and ethnicity data has improved over time since the introduction of the T-MSIS system; however, there is still a long way to go. We encourage you to check out the quality of T-MSIS data in your state and advocate for better data quality. CMS needs to continue to push for and provide support for improvements in data collection and reporting. In addition, CMS should avoid using imputed race and ethnicity data in their own reports as this practice masks and perpetuates data quality issues. Ultimately, CMS, states, health plans, and advocates need to work together to improve the quality of race and ethnicity data among Medicaid beneficiaries in T-MSIS.

When is the appropriate time to collect race and ethnicity data? MACPAC advises collecting race and ethnicity data during the Medicaid application or redetermination process. To improve health equity data, CMS recommends promoting the utilization of existing and new methods of collecting demographic data and collaborating with partners, such as the Social Security Administration, to collect race and ethnicity data. For the Medicaid population, we need to ensure that this data is being incorporated into T-MSIS.

While this is an opportune time to collect data, it must be made clear why this data is being collected and how it will be used. Among adults receiving Medicaid LTSS, there are

several additional opportunities to collect race and ethnicity data beyond the application process. For example, everyone receiving HCBS is required to go through a person-centered planning process. This is an ideal time to collect race and ethnicity data as the individual's support team can ask about their race and ethnicity while also explaining the purpose of collecting this information. For individuals receiving institutional LTSS, nursing homes and other institutional facilities should be collecting this data. While plans and providers may be collecting this data, it is critical that this data be reported to states, and states must input this information into T-MSIS.

While plans and providers may be collecting this data, ***it is critical that this data be reported to states, and states must input this information into T-MSIS.***

Funding Statement

Funding for this brief was provided by the National Institute on Disability, Independent Living, and Rehabilitation Research, grant numbers 90RTCP0006 and 90RTCP0009. NIDILRR is a center within the Administration for Community Living (ACL), U.S. Department of Health and Human Services (HHS). This brief 's contents reflect the views of the authors and not of the National Institute on Disability, Independent Living, and Rehabilitation Research.

Sources

Medicaid.gov. *Transformed Medicaid Statistical Information System (T-MSIS).* https://www.medicaid.gov/medicaid/data-systems/macbis/transformed-medicaid-statistical-information-system-t-msis/index.html

DQ Atlas. *DQ Assessments: Race and Ethnicity (2021).* https://www.medicaid.gov/dq-atlas/landing/topics/single/map?topic=g3m16&tafVersionId=35

Lines, L. M., Humphrey, J. L., & Barch, D. H. (2022). Imputing race and ethnicity: A fresh voices commentary from the Medical Care Blog. *Medical Care* 60(5), 351-356, May 2022. | DOI: 10.1097/MLR.0000000000001717 http://doi.org/10.1097/MLR.0000000000001717

Randall, M., Stern, A., & Su Y. (2021). *Five Ethical Risks to Consider before Filling Missing Race and Ethnicity Data.* Urban Institute. https://www.urban.org/sites/default/files/publication/103830/five-ethical-risks-to-consider-before-filling-missing-race-and-ethnicity-data-workshop-findings.pdf

Medicaid.gov. *Reports & Evaluations: Long Term Services and Supports Users and Expenditures.* https://www.medicaid.gov/medicaid/long-term-services-supports/reports-evaluations/index.html

Carpenter, A., Eckstein, M., Murray, C., Stepanczuk, C., & Wysocki, A. *Data Notes for Medicaid TAF Long-Term Services and Supports Annual Expenditures and Users, 2019-2021.* Mathematica. https://www.medicaid.gov/medicaid/long-term-services-supports/downloads/ltss-data-notes-taf-users-expenditures.pdf

MACPAC. (2023). *Medicaid Race and Ethnicity Data Collection and Reporting: Recommendations for Improvement.* Report to Congress on Medicaid and CHIP. https://www.macpac.gov/wp-content/uploads/2023/03/Chapter-1-Medicaid-Race-and-Ethnicity-Data-Collection-and-Reporting-Recommendations.pdf

CMS. (2022). *The Path Forward: Improving Data to Advance Health Equity Solutions.* Office of Minority Health. https://www.cms.gov/files/document/path-forwardhe-data-paper.pdf

Appendix

Table 1. Percent of Adults who are Missing Race and Ethnicity Data by State and LTSS Status (Institutional LTSS, HCBS, or Non-LTSS)

State	Institutional LTSS	HCBS	Non-LTSS
AK	6.31%	5.80%	6.89%
AR	4.09%	2.67%	12.21%
AZ	23.81%	22.87%	18.32%
CA	9.31%	8.08%	10.08%
CO	0.34%	0.39%	0.66%
CT	9.06%	10.25%	25.31%
DC	55.50%	34.06%	25.68%
DE	0.00%	0.00%	0.00%
FL	13.89%	22.77%	14.63%
GA	4.30%	16.46%	10.30%
HI	26.09%	16.58%	19.60%
IA	54.11%	27.54%	21.36%
ID	0.85%	0.09%	0.59%
IL	6.93%	3.90%	6.04%
IN	10.88%	22.18%	17.06%
KS	13.97%	3.34%	5.04%
KY	18.20%	26.68%	18.23%
LA	13.76%	3.76%	9.52%
MA	65.31%	45.66%	44.14%
MD	7.25%	13.30%	17.54%
ME	4.58%	2.43%	6.95%
MI	5.62%	2.51%	3.69%
MN	3.50%	2.33%	8.40%

State	Institutional LTSS	HCBS	Non-LTSS
MO	6.41%	5.03%	5.60%
MT	17.52%	5.60%	9.91%
NC	0.64%	0.61%	0.75%
ND	0.00%	0.00%	0.92%
NE	6.70%	3.66%	5.63%
NH	11.65%	4.28%	10.07%
NJ	11.00%	20.12%	10.14%
NM	1.29%	0.80%	1.14%
NV	2.07%	1.76%	3.01%
NY	8.29%	8.28%	31.83%
OH	10.19%	4.17%	6.45%
OK	0.87%	0.30%	5.32%
OR	7.20%	17.71%	20.53%
PA	5.44%	4.18%	6.41%
RI	0.00%	0.00%	0.35%
SC	15.14%	28.90%	33.02%
SD	0.00%	0.00%	0.00%
TN	7.07%	7.06%	47.67%
TX	14.36%	16.00%	9.74%
UT	45.41%	26.47%	44.03%
VA	0.26%	0.99%	7.30%
VT	39.70%	18.13%	13.67%
WA	6.54%	6.12%	4.99%
WI	8.91%	9.70%	6.32%
WV	10.02%	13.60%	13.64%
WY	49.83%	18.90%	19.32%