# Ensuring Quality in TAF Research: Introducing a Reporting Checklist

February 29, 2024

With support from the Commonwealth Fund and the Robert Wood Johnson Foundation

# Housekeeping

- Please use the **Q&A function** to ask questions throughout the presentations. We will pull from the submitted questions during the Q&A portion.

- This webinar is being **recorded** and will be made available on AcademyHealth's website following the event. You will receive an email when the recording is available.

# Agenda

1. Welcome & Introductions
2. CMS Data Policy Change
3. A New Resource: The TAR Checklist
4. Reflections on the Utility and Importance of a TAF Checklist
5. Application of the Checklist to a Research Project
6. Q&A

AcademyHealth

# CMS Data Policy Change

**William Schpero, PhD,** *Weill Cornell Medical College*

# MDLN CMS RFI Response

- CMS announced it is **discontinuing access to physical data** extracts beginning in August 2024.
- **AcademyHealth** issued a response on February 26.
- **MDLN** is drafting a response by March 29 RFI deadline.
- We appreciate your insights, especially from Medicaid researchers who are familiar with the VRDC.
- **Email** Annaliese.Johnson@academyhealth.org with comments.

# A New Resource: The TAR Checklist

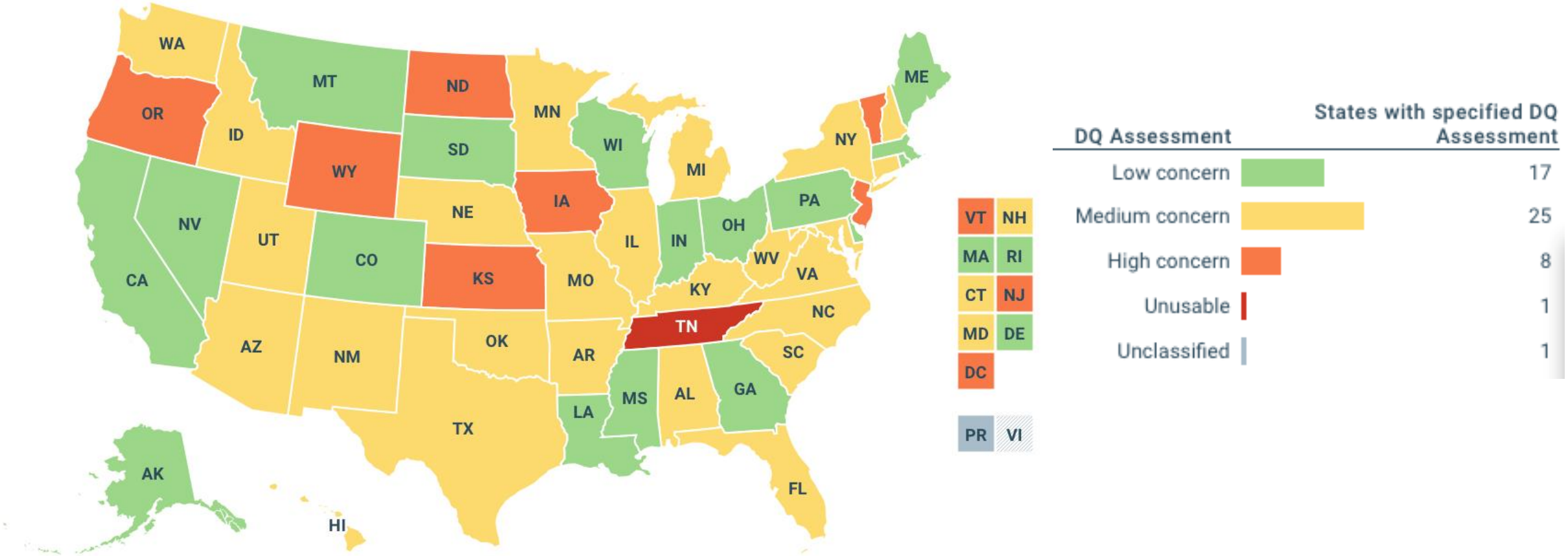
**William Schpero, PhD,** *Weill Cornell Medical College*

# Background: Motivation

- The **T-MSIS Analytic Files** (TAF) are an important resource for conducting timely, policy-relevant research on the Medicaid and CHIP programs.
- TAF data are also **highly complex**, with varying quality across data elements, states, and time.
- In recognition of these challenges, the MDLN has drafted a **TAF Analysis Reporting (TAR) Checklist** as a guide for those who generate and evaluate analyses of the TAF national Medicaid enrollment and claims data.

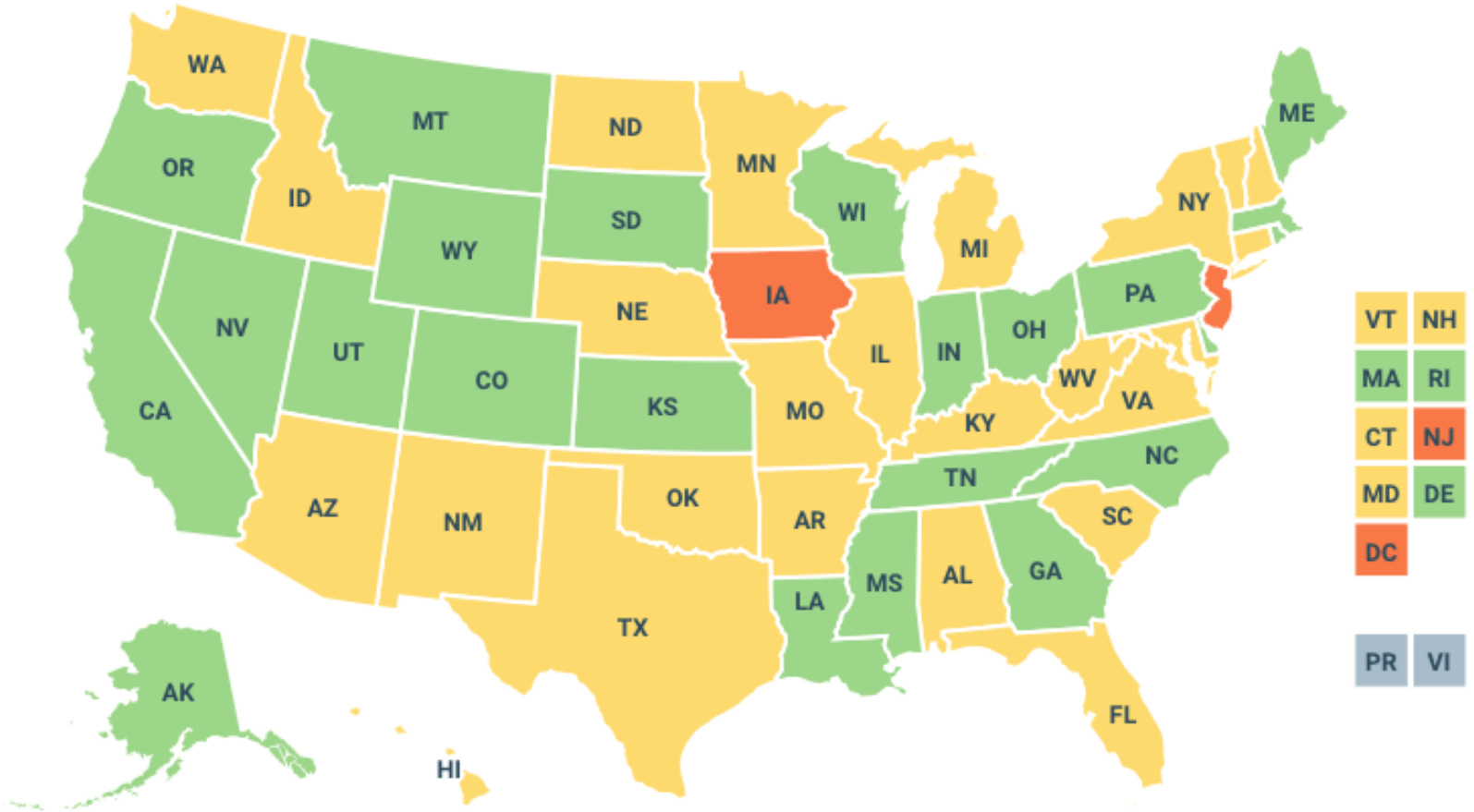
# Background: Motivation

**Eligibility Group Code Quality (DQ Atlas, 2016 Release 2)**



# Background: Motivation

**Eligibility Group Code Quality (DQ Atlas, 2021 Release 1)**




| DQ Assessment | States with specified DQ Assessment |
|---|---|
| Low concern | 23 |
| Medium concern | 25 |
| High concern | 3 |
| Unusable | 0 |
| Unclassified | 2 |

# Background: Process

- Initial draft developed by MDLN sub-committee
- Full MDLN provided input
- Revised draft shared with external stakeholders
  - MDLN Advisory Group
  - Health policy journal editors
- **Today**: Draft shared publicly for feedback
- **Soon**: Revised draft to be submitted for peer review

# Background: Format

**Checklist**

| Category | Description | Examples |
|---|---|---|
| **Data** | | |
| Files, Years, and Release Versions | (a) Indicate which TAF files were used in the analysis (e.g., Demographic and Eligibility, Inpatient, Other Services, etc.).<br>(b) Indicate whether the study drew from 100% TAF files or pre-specified extracts.<br>(c) Indicate which years of TAF data were included in the analysis.<br>(d) Indicate which file versions were included in the analysis (e.g., preliminary, release 1, release 2, etc.). | "Data from this study came from the TAF Other Services and Annual Provider Files for 2017 (Release 2), 2018 (Release 2), and 2019 (Release 1).<br><br>"For this analysis, we used Release 1 of the Inpatient claims file (IP) for data years 2019 and 2020."<br><br>"Please see Appendix Table A1 for a summary of which files and TAF release versions were used by year."<br><br>"We generated our study cohort from a 100% TAF sample for 2017-2018 Release 2."<br><br>"We derived our study cohort from a 2016 Release 1 20% sample of Medicaid beneficiaries between the ages of 19-44 in Texas, California, and New York." |
| | | |
| **Cohort** | | |
| Eligibility Criteria | (a) If applicable, describe what eligibility category codes were used to identify the study sample and whether they were | "We limited our analysis to individuals newly eligible under the Affordable Care Act's Medicaid expansion as identified by an |

# Part 1: Data Details

**Files, Years, and Release Versions**

a) Indicate which TAF files were used in the analysis (e.g., Demographic and Eligibility, Inpatient, Other Services, etc.).
b) Indicate whether the study drew from 100% TAF files or pre-specified extracts.
c) Indicate which years of TAF data were included in the analysis.
d) Indicate which file versions were included in the analysis (e.g., preliminary, release 1, release 2, etc.).

# Part 2: Cohort

**Eligibility Criteria**

a) If applicable, describe what eligibility category codes were used to identify the study sample and whether they were used in combination with any other variables (e.g., age, receipt of specific medical services, etc.).

**Enrollment Span**

a) If applicable, indicate the minimum period of enrollment required for an enrollee to be included in the study sample and how the enrollment period was defined.

# Part 2: Cohort

**Scope of Benefits**

a) Indicate whether the analysis included enrollees with full scope, comprehensive, or restricted benefits.

**Dual Eligibility**

a) Describe whether individuals dually enrolled in Medicare and Medicaid were included in or excluded from the study sample and, if applicable, how dual eligibility was defined.

# Part 3: State and Territory Exclusions

**Criteria**

a) Indicate which states and/or territories were included (or excluded) from the analysis on the basis of data quality concerns.
b) Indicate the criteria by which state exclusions were made, including measures, data sources, and thresholds.

**State variation table**

a) Include a state-level table (which may appear in an appendix) summarizing the number of observations, means, medians, and missingness for key study measures.

# Part 4: Special Considerations

**Encounter Data**

a) Indicate whether the analysis excluded either fee-for-service or managed care enrollees. If managed care enrollees were excluded, define the criteria used to do so.

b) Indicate which types of claims records (e.g., fee-for-service claims, service tracking claims, capitation payments, etc., see variable CLM_TYPE_CD) were included in the analysis.

# Part 4: Special Considerations

**Spending**

a) Indicate which types of claims records (e.g., fee-for-service claims, service tracking claims, capitation payments, etc., see variable CLM_TYPE_CD) were included to measure spending.

b) If including service-specific spending for managed care encounters, indicate how spending was imputed (payments from plans to providers on encounter records are generally redacted).

# Part 4: Special Considerations

**Using TAF with Predecessor Medicaid Analytic eXtract (MAX) Data**

a) Indicate if the analysis included data from the Medicaid Analytic eXtract (MAX) and, if so, for what years and which states.
b) If applicable, include an exhibit examining trends in key measures by state over time and particularly during any transition from MAX to TAF.

# Notes

- Checklist items are **recommendations**, not prescriptions, and can be adapted to different study designs.

- **Not all** checklist items will apply to all studies.

- The TAR Checklist is **not meant to be a substitute** for broader reporting checklists (e.g., the STROBE checklists for observational research).

- This checklist is a **living document**; it may be updated over time as the TAF data evolve.

AcademyHealth

# Reflections on the Utility and Importance of a TAF Checklist

**Julie Donohue, PhD,** *University of Pittsburgh*

# Reflections on the Utility and Importance of a TAF Checklist

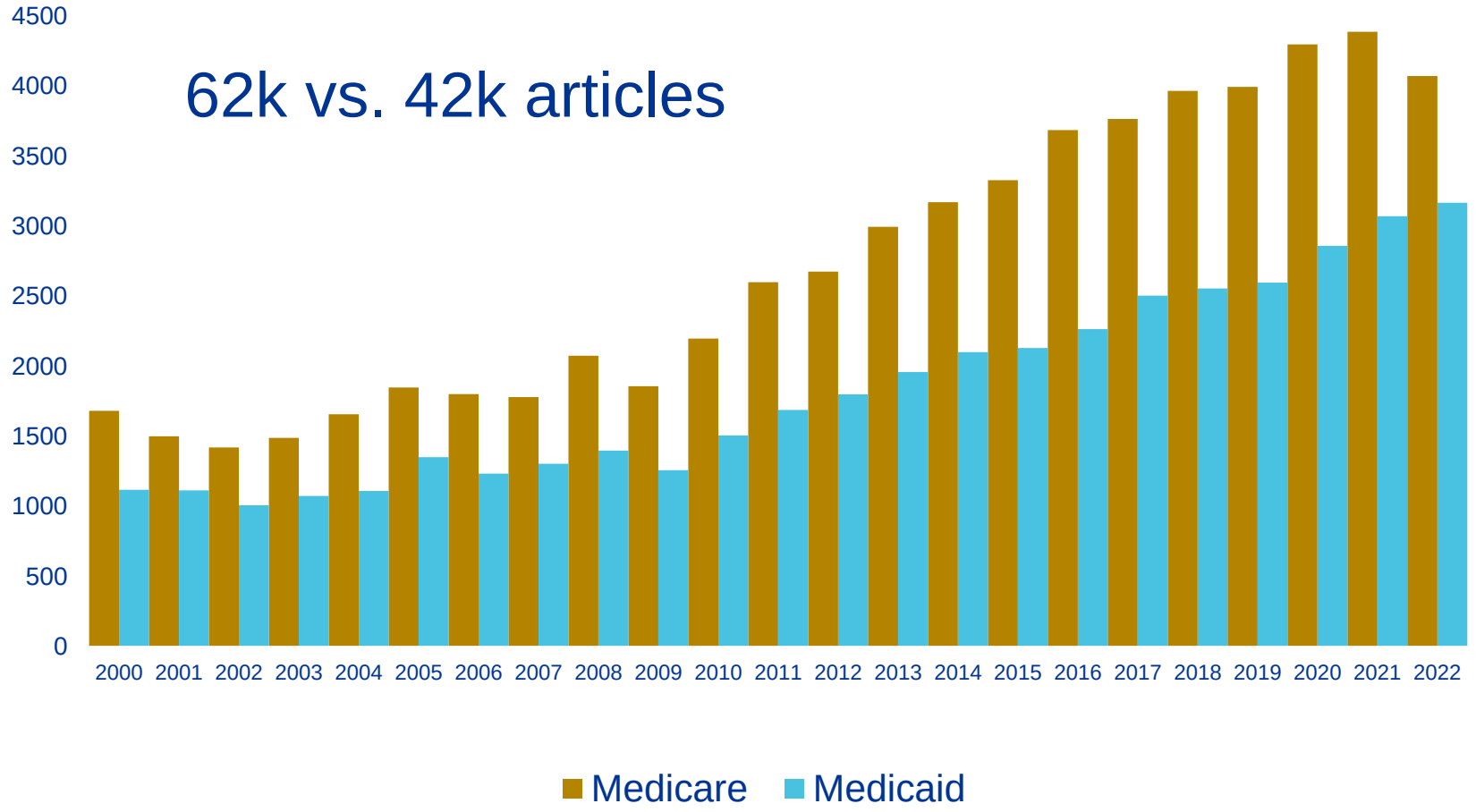
Julie Donohue, PhD

University of Pittsburgh

University of Pittsburgh | Health Policy and Management | School of Public Health

# Medicare vs. Medicaid

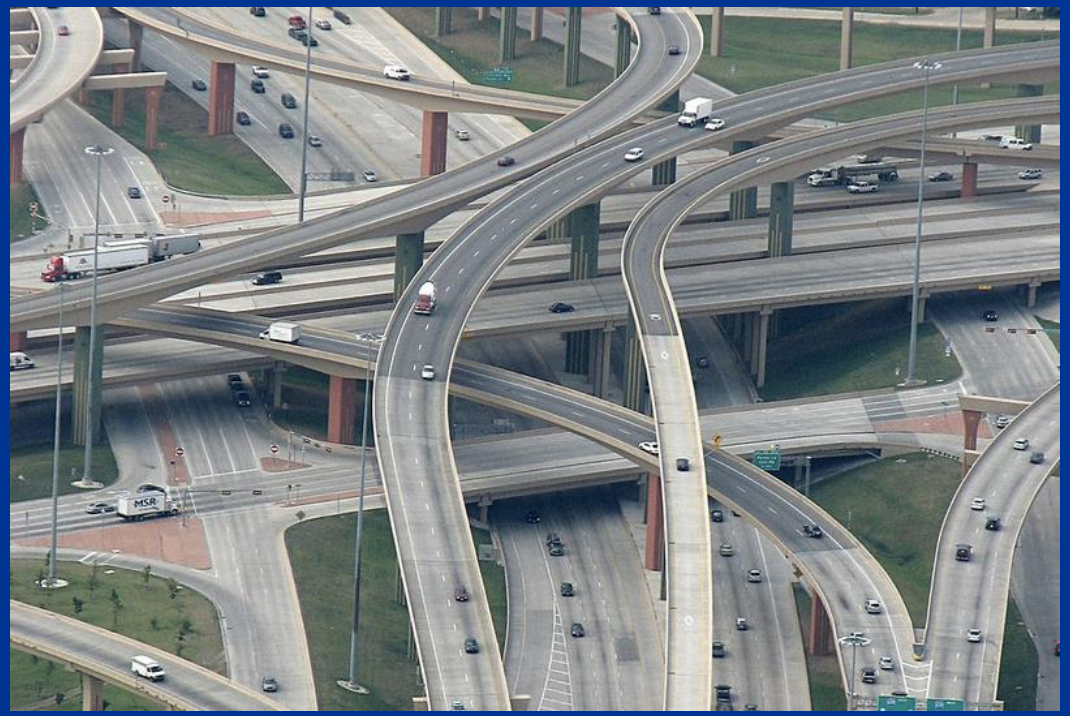
- 2023 enrollment
  - 66m in Medicare
  - 87m in Medicaid



# T-MSIS Analytic Files

- Growing community of researchers working with Medicaid
  - Background conducting Medicaid research in other data
  - Background working with other claims data
  - Background in neither
- ALL need training in how to work with TAF

# TAR Checklist



# Helps orient researchers to Medicaid / TAF

- 50+ Medicaid programs
- Each state-year could differ
  - Populations covered
  - Benefits/services/payment >>all affect measurement
  - Data quality and completeness
- Managed care vs. FFS distinction differs from Medicare
- Enrollment dynamics differ from other payers
- Multiple eligibility groups with different benefits

# How a checklist helps editors

- Helps editors to learn about new data
  - what they should ask authors to do
- Improves validity of studies
- Facilitates transparency and reproducibility
- Facilitates comparisons across studies
- Improves impact if policymakers have confidence researchers know what they are doing with their data

AcademyHealth

# Application of the Checklist to a Research Project

**Laura Barrie Smith, PhD,** *Urban Institute*

Medicaid Data Learning Network (MDLN) Webinar

# Using the TAF Reporting Checklist as a Researcher/Author

**February 29, 2024**

URBAN INSTITUTE · ELEVATE · THE · DEBATE

Laura Barrie Smith, PhD

# When to use the TAR checklist

1. Before starting your study
   - To inform research questions and feasibility
   - To select file types (DE, IP, etc.) and years to request from CMS
2. While finalizing study design
   - To select state(s) and define study population(s)
   - To specify variables
3. Throughout analysis
   - To create all necessary tables and output
   - To address unanticipated issues
4. While writing / just before submitting to a journal

# Using the TAR: Example

- Research objective: to examine access to recommended long-term asthma control medications among Medicaid/CHIP-enrolled children with asthma, overall and by race and ethnicity
- Study population: Medicaid-enrolled children with asthma
- Study period: 2019, with lookback in 2018 to identify previous asthma diagnoses
- Key outcome variable: asthma prescription medication fills

# Data: Files, Years, and Release Versions

| Category | Description |
|---|---|
| **Data** | |
| Files, Years, and Release Versions | (a) Indicate which TAF files were used in the analysis (e.g., Demographic and Eligibility, Inpatient, Other Services, etc.).<br>(b) Indicate whether the study drew from 100% TAF files or pre-specified extracts.<br>(c) Indicate which years of TAF data were included in the analysis.<br>(d) Indicate which file versions were included in the analysis (e.g., preliminary, release 1, release 2, etc.). |

**(a) DE, RX, IP, OT**
**(b) 100% TAF**
**(c) 2018-2019**
**(d) 2018 Release 2; 2019 Release 1**

# Cohort: Eligibility Criteria

| Cohort | |
|---|---|
| Eligibility Criteria | (a) If applicable, describe what eligibility category codes were used to identify the study sample and whether they were used in combination with any other variables (e.g., age, receipt of specific medical services, etc.). |

**(a) No eligibility category restrictions; include children ages 2-18 (defined according to age as of January 1, 2019)**

# Cohort: Enrollment Span

| | |
|---|---|
| Enrollment Span | (a) If applicable, indicate the minimum period of enrollment required for an enrollee to be included in the study sample and how the enrollment period was defined. |

**(a) Require continuous enrollment for CY2018-2019**
- **(i) defined as at least 330 days during each year**
- **(ii) sensitivity analysis relaxed this requirement**

# Cohort: Scope of Benefits

| | |
|---|---|
| Scope of Benefits | (a) Indicate whether the analysis included enrollees with full scope, comprehensive, or restricted benefits. |

**(a) Exclude enrollees with restricted benefits any month 2018-2019**

# Cohort: Dual Eligibility

| | |
|---|---|
| Dual Eligibility | (a) Describe whether individuals dually enrolled in Medicare and Medicaid were included in or excluded from the study sample and, if applicable, how dual eligibility was defined. |

**(a) Exclude those dually enrolled any month 2018-2019**

Appendix Table AX: Definition of Study Cohort

| Exclusion | Number of enrollees remaining | Number Excluded from Last Step | Share Excluded from Last Step | | |
|---|---|---|---|---|---|
| | | | Overall | State Lower Range | State Upper Range |
| All enrollee records in the Demographics and Eligibility TAF file in our states with nonmissing birth date and enrollee ID, 2019 | 33,101,866 | 0 | 0.0% | 0.0% | 0.0% |
| Exclude enrollees 1 or younger or 19 or older as of Jan 1, 2019 | 12,229,464 | 20,872,402 | 63.1% | 45.4% | 68.9% |
| Exclude enrollees with missing eligibility information or sex or missing/invalid ZIP code | 12,195,310 | 34,154 | 0.3% | 0.0% | 0.6% |
| Exclude enrollees not continuously enrolled in Medicaid or CHIP for at least 330 days in 2019 | 9,341,958 | 2,853,352 | 23.4% | 16.2% | 37.4% |
| Exclude enrollees dually eligible for Medicare in any month of 2019 | 9,340,434 | 1,524 | 0.0% | 0.0% | 0.1% |
| Exclude enrollees with restricted Medicaid benefits | 9,239,640 | 100,794 | 1.1% | 0.0% | 5.4% |
| Exclude enrollees with duplicate eligibility records | 9,238,839 | 801 | 0.0% | 0.0% | 0.1% |
| Exclude enrollees who we do not identify as Black NH or white NH | 3,774,489 | 5,464,350 | 59.1% | 7.9% | 78.7% |
| Exclude enrollees not continuously enrolled in Medicaid or CHIP for at least 330 days in 2018 | 3,284,057 | 490,432 | 13.0% | 8.4% | 20.1% |
| Exclude enrollees without a 2018 asthma diagnosis | **272,740** | 3,011,317 | 91.7% | 89.3% | 95.6% |

# State and Territory Exclusions: Criteria

| State and Territory Exclusions | |
|---|---|
| Criteria | (a) Indicate which states and/or territories were included (or excluded) from the analysis on the basis of data quality concerns.<br>(b) Indicate the criteria by which state exclusions were made, including measures, data sources, and thresholds. |

**(a) Using DQ Atlas, require "low concern" for key variables (bene age, ZIP, total Medicaid/CHIP enrollment, RX claims volume); require "low concern" or "medium concern" for secondary variables: IP claims volume, OT claims volume**

**(b) Using our own assessment of data quality, exclude states where the share of Black, non-Hispanic or white, non-Hispanic kids differs by more than 10 percentage points from ACS benchmarks or where > 20% of benes have a missing value**

Appendix Table AX: Selection of States for Analysis

| Criterion | Remaining States | Count of remaining states |
|---|---|---|
| All states | AK, AL, AZ, AR, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, SC, SD, TN, TX, UT, VA, VT, WA, WI, WV, and WY | 51 |
| "Low Concern" on beneficiary age in 2018 and 2019 | AK, AL, AZ, AR, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, SC, SD, TN, TX, UT, VA, VT, WA, WI, WV, and WY | 51 |
| "Low Concern" on Total Medicaid and CHIP Enrollment in 2018 and 2019 | AK, AL, AZ, AR, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS, LA, MA, MD, MI, MN, MO, MS, MT, NC, ND, NE, NH, NM, NV, NY, OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, VT, WA, WI, and WY | 46 |
| "Low Concern" on beneficiary ZIP code in 2018 and 2019 | AK, AL, AZ, AR, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS, LA, MA, MD, MI, MN, MO, MS, MT, NC, ND, NE, NH, NM, NV, NY, OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, WA, WI, and WY | 45 |
| "Low Concern" on RX claims volume in 2019 | AZ, CA, CT, DC, DE, GA, HI, IA, ID, IL, IN, KS, LA, MA, MD, MI, MO, MS, MT, ND, NE, NH, NM, NV, NY, OR, PA, TN, TX, UT, VA, WA, and WI | 33 |
| Exclude states with "high concern" or "unusable" IP claims volume in 2018 | AZ, CA, DC, DE, HI, IA, ID, IL, IN, KS, LA, MD, MI, MO, MS, MT, ND, NE, NM, NV, OR, PA, TN, TX, UT, VA, WA, and WI | 28 |
| Exclude states with "high concern" or "unusable" OT claims volume in 2018 | AZ, CA, DC, DE, HI, IA, ID, IL, IN, KS, LA, MD, MI, MO, MS, MT, ND, NE, NM, NV, OR, PA, TN, TX, UT, VA, WA, and WI | 28 |
| Able to identify Black, non-Hispanic and white, non-Hispanic enrollees based on data quality assessment of race and ethnicity variable, 2019 | CA, DE, IN, MI, MS, NE, NM, NV, PA, TX, and WA | 11 |
| For sub-analysis examining the relationship with preventable hospitalization | | |
| Exclude states with "high concern" or "unusable" IP or OT claims volume in 2019 | CA, DE, IN, MI, MS, NE, NM, NV, PA, TX, and WA | 11 |

# State and Territory Exclusions: State variation table

| | |
|---|---|
| State variation table | (a) Include a state-level table (which may appear in an appendix) summarizing the number of observations, means, medians, and missingness for key study measures. |

**(a) Run all steps of analysis by state to identify state-specific issues or outliers. Provide key tables by state in the manuscript body or Appendix.**

## II.i Sample Characteristics by State

*Table A3: By State: Preventable Hospitalizations per 100,000 Pediatric Medicaid Enrollees by Enrollee Characteristics and Extent of ZIP Code-Level Racialized Economic Segregation, 2018*

*Panel A: Alabama, Alaska, Arizona, and Arkansas*

| | Alabama | | Alaska | | Arizona | | Arkansas | |
|---|---|---|---|---|---|---|---|---|
| | Composite Measure | Share of enrollees | Composite Measure | Share of enrollees | Composite Measure | Share of enrollees | Composite Measure | Share of enrollees |
| Overall | 109 | 100% | 81 | 100% | 164 | 100% | 109 | 100% |
| *Sex* | | | | | | | | |
| Female | 101 | 49% | 95 | 49% | 167 | 49% | 122 | 49% |
| Male | 116 | 51% | 68 | 51% | 161 | 51% | 97 | 51% |
| *Age in Years* | | | | | | | | |
| 0-5 | N/A | 36% | N/A | 36% | N/A | 35% | N/A | 35% |
| 6-9 | 128 | 23% | 70 | 23% | 209 | 22% | 131 | 23% |
| 10-14 | 88 | 27% | 84 | 27% | 141 | 28% | 92 | 28% |
| 15-17 | 117 | 14% | 92 | 14% | 139 | 15% | 107 | 14% |

Source: Smith LB, O'Brien C, Kenney GM, Tabb LP, Verdeflor A, Wei K, Lynch V, Waidmann T. "Racialized Economic Segregation and Potentially Preventable Hospitalizations among Medicaid/CHIP-enrolled children" *Health Services Research*. 2023; 58(3):599-611. https:// doi.org/10.1111/1475-6773.14120

# Special considerations: Encounter Data

| Special Considerations | |
|---|---|
| Encounter Data | (a) Indicate whether the analysis excluded either fee-for-service or managed care enrollees. If managed care enrollees were excluded, define the criteria used to do so.<br>(b) Indicate which types of claims records (e.g., fee-for-service claims, service tracking claims, capitation payments, etc., see variable CLM_TYPE_CD) were included in the analysis. |

**(a) Include both FFS and managed care enrollees**
**(b) Restrict to CLM_TYPE_CD in (1,3,A,C)**

# Special considerations: Spending

| | |
|---|---|
| Spending | (a) Indicate which types of claims records (e.g., fee-for-service claims, service tracking claims, capitation payments, etc., see variable CLM_TYPE_CD) were included to measure spending.<br>(b) If including service-specific spending for managed care encounters, indicate how spending was imputed (payments from plans to providers on encounter records are generally redacted). |

N/A

# Special considerations: Using TAF with Predecessor Medicaid Analytic eXtract (MAX) Data

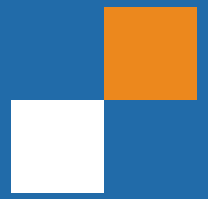
| | |
|---|---|
| Using TAF with Predecessor Medicaid Analytic eXtract (MAX) Data | (a) Indicate if the analysis included data from the Medicaid Analytic eXtract (MAX) and, if so, for what years and which states.<br>(b) If applicable, include an exhibit examining trends in key measures by state over time and particularly during any transition from MAX to TAF. |

N/A

# Examples

- Smith LB, O'Brien C, Kenney GM, Tabb LP, Verdeflor A, Wei K, Lynch V, Waidmann T. "Racialized Economic Segregation and Potentially Preventable Hospitalizations among Medicaid/CHIP-enrolled children" *Health Services Research*. 2023; 58(3):599-611. https:// doi.org/10.1111/1475-6773.14120
- O'Brien C, Smith LB, Waidmann T, Kenney GM. "Preventable Hospitalizations among Adult Medicaid Enrollees in 2019: Variation by Supplemental Security Income Eligibility Status, Race, and State" Washington DC: Urban Institute, January 2024. https://www.urban.org/research/publication/preventable-hospitalizations-among-adult-medicaid-enrollees-2019
- Smith LB, O'Brien C, Kenney GM, Waidmann TA. "Black-White Disparities in Asthma Hospitalizations and ED Visits Among Medicaid-Enrolled Children" *Hospital Pediatrics*. 2024; forthcoming.
- Smith LB, O'Brien C, Kenney GM. "Examining Race and Ethnicity Data Quality for Medicaid/CHIP-Enrolled Children in the T-MSIS Analytic Files: A State-by-State Resource for Researchers." Washington DC: Urban Institute, March 2023. https://www.urban.org/research/publication/examining-race-and-ethnicity-data-quality-medicaidchip-enrolled-children-tmsis-analytic-files

Contact: LaSmith@urban.org



# Q&A

AcademyHealth

# Thank you!

If you have any input you would like to share on the **T-MSIS Analytic Files Analysis Reporting (TAR) Checklist**, please reach out:

- William Schpero: wls4001@med.cornell.edu

As always, we welcome any questions or other comments:

- Sarah Gordon: gordonsh@bu.edu
- John McConnell: mcconnjo@ohsu.edu
- Susan Kennedy: susan.kennedy@academyhealth.org