

Diversifying the FDA's Sentinel System with Rigorous Quality Inclusion Rules for the U.S. Medicaid Population

Michael Stagner¹, Judith C. Maro², Sarah K. Dutcher³, David Moeny³, Robert Rosofsky⁴, Daniel Kiernan², Laura Shockro², Alexander Mai², Jessica Pritchard¹, Steve Lippmann¹, Pratap Adhikari¹, Bradley G. Hammill¹

¹ Department of Population Health Sciences, Duke University School of Medicine

² Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute

³ Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration

⁴ Health Information Systems Consulting

Background The U.S. Medicaid and Children's Health Insurance Programs (CHIP) insure over 25% of the population and about 40-50% of pregnancies following expansions to the program over the last decade. Adding this low-income population to the U.S. FDA's Sentinel System is a high priority, but data quality variability among the 50+ contributing jurisdictions requires rigorous preliminary review.

Objectives To establish jurisdiction-level, beneficiary-level, and record-level criteria for inclusion of Medicaid/CHIP data into a Sentinel Common Data Model (SCDM)-compliant database and to document the fit-for-purpose requirements of the Medicaid/CHIP data for Sentinel System regulatory needs.

Methods Transformed Medicaid Statistical Information System (T-MSIS) analytic files (TAF) from 2014 to 2018 were utilized. We prioritized TAF variables relevant to the needs of drug safety analyses and identified jurisdictions (states, territories, and the District of Columbia) with acceptable quality on these variables by year and by plan type as reported by Medicaid's Data Quality (DQ) Atlas. Based on TAF documentation, we devised beneficiary-level inclusion rules to identify individuals for whom Medicaid/CHIP was the primary insurer; and established record-level criteria to exclude records related only to administrative payments.

Results We selected 12 high priority variables, from over 80 available, within the DQ Atlas to assess fit-for-purpose data for jurisdiction-year-plan combinations with an acceptable level of quality. These variables covered different topics but prioritized complete capture of healthcare utilization and included enrollment, eligibility, claims file completeness, and service use. In 2018, data from 44 jurisdictions met these standards (9 of which met these standards for all 5 years of data included), while 9 jurisdictions were excluded because at least one of the selected high priority variables had poor data quality. We did not exclude source data based on the quality of demographic or provider information since data deficiencies in these domains can be handled during data analysis. Data from about 20% of beneficiaries were excluded due to either dual Medicare/Medicaid eligibility (since Medicare is the primary payer) or eligibility for only partial benefits. ~45% of records from the "other services" file were identified as capitated payment records and were excluded from data transformation.

Conclusions The U.S. Medicaid/CHIP data is a rich resource of variable quality due to its aggregation across 50+ jurisdictions with differing rules and standards. A rigorous process to determine initial fit-for-purpose criteria among these different data sources is needed as a formal step prior to routine data characterization and quality review to optimize resource management.

Disclosures Dr. Hammill reports a research grant from iRhythm Technologies. All other authors report nothing to disclose.

Funding This work was supported by the Office of the Secretary PCORTF under Interagency Agreement #750121PE080007 Indefinite Delivery/Indefinite Quantity (IDIQ) Sentinel Contract 75F40119D10037 Task Order 75F40119F19001 from the US Food and Drug Administration (FDA) and the Assistant Secretary for Planning and Evaluation (ASPE) within the U.S. Department of Health and Human Services. The views expressed in this presentation represent those of the presenters and do not necessarily represent the official views of the U.S. FDA or ASPE.

The U.S. Medicaid/CHIP data is a rich resource of **variable quality** due to its aggregation across 50+ jurisdictions with **differing rules and standards**. A rigorous process to determine **fit-for-purpose criteria** among these different data sources is needed as a formal step **prior to use** for generating real-world evidence.

DQ Atlas Categories & Topics Used for Data Quality Assessment in the Data Transformation for Sentinel

Jurisdiction-plans having data considered unusable on any of the topics below are excluded from data transformation.

For detailed descriptions of each topic, see the DQ Atlas, <https://www.medicaid.gov/dq-atlas>.

Enrollment over time

- Number of enrollment spans
- Length of enrollment gaps
- Overlapping enrollment spans, %

Beneficiary eligibility

- Dual eligibility codes, % missing
- Restricted benefits codes, % missing

Claim file completeness

- Service users with any utilization, %
- Claims volume per 1000 enrolled months
- CMC plan encounters per 1000 CMC-enrolled months

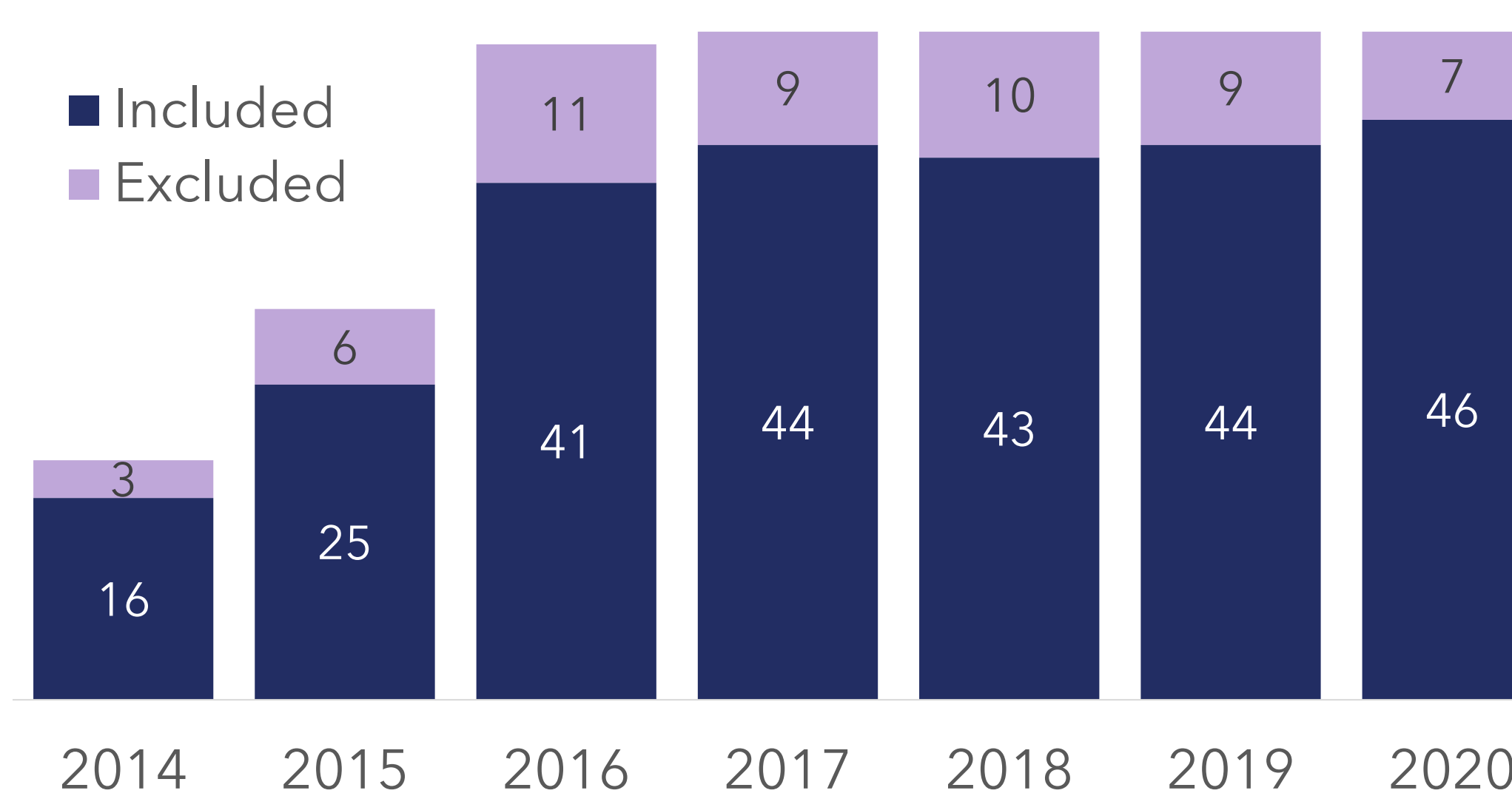
Service use information

- Admission dates, % missing
- Discharge dates, % missing
- Diagnosis codes, % with any valid
- Types of service codes, % missing
- Procedure codes, % with any valid
- National drug codes, % missing

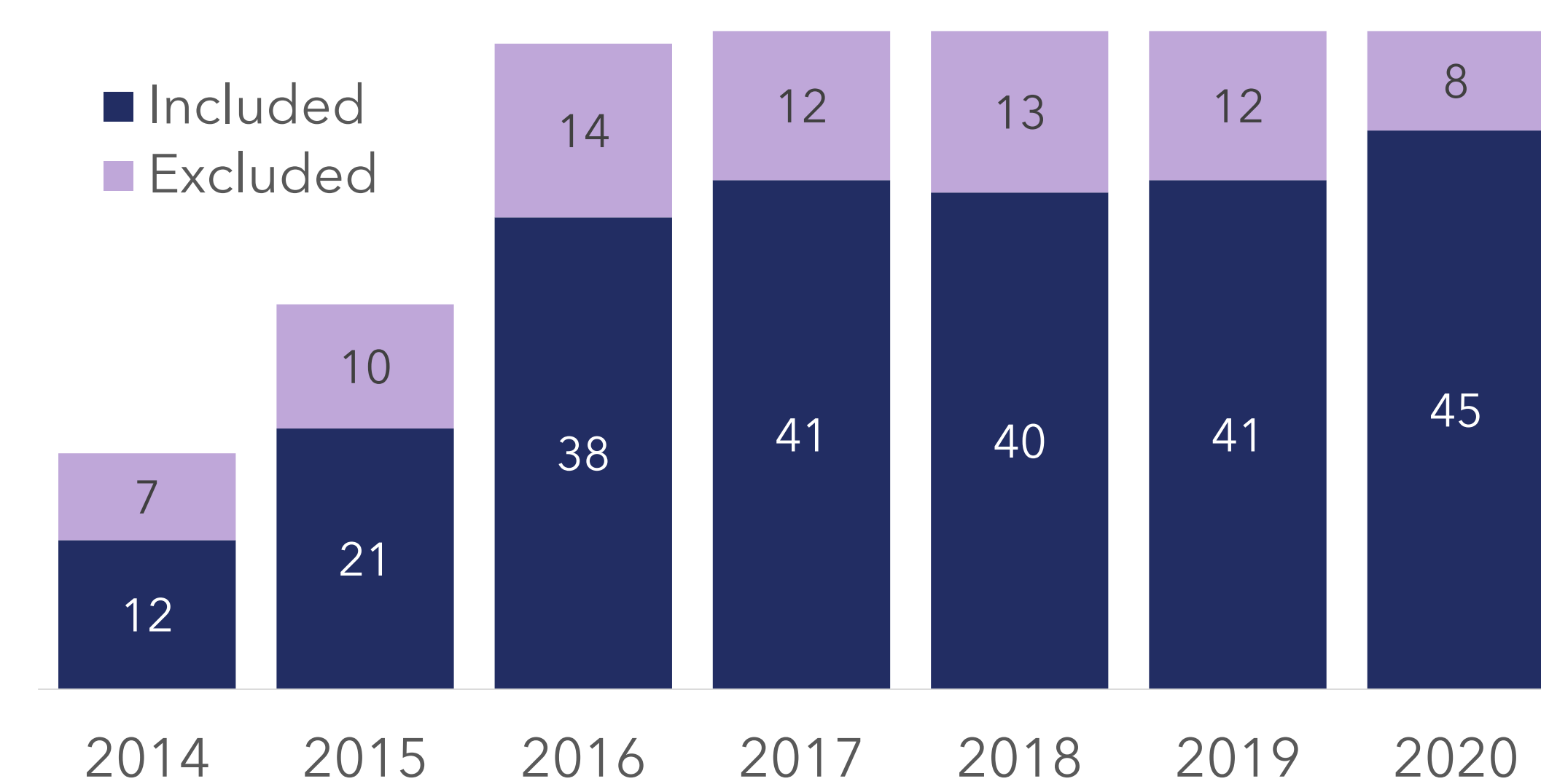
Number of Jurisdictions Included/Excluded By Year After Applying Data Quality Criteria, 2014-2020

The 53 jurisdictions considered include the 50 U.S. states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands

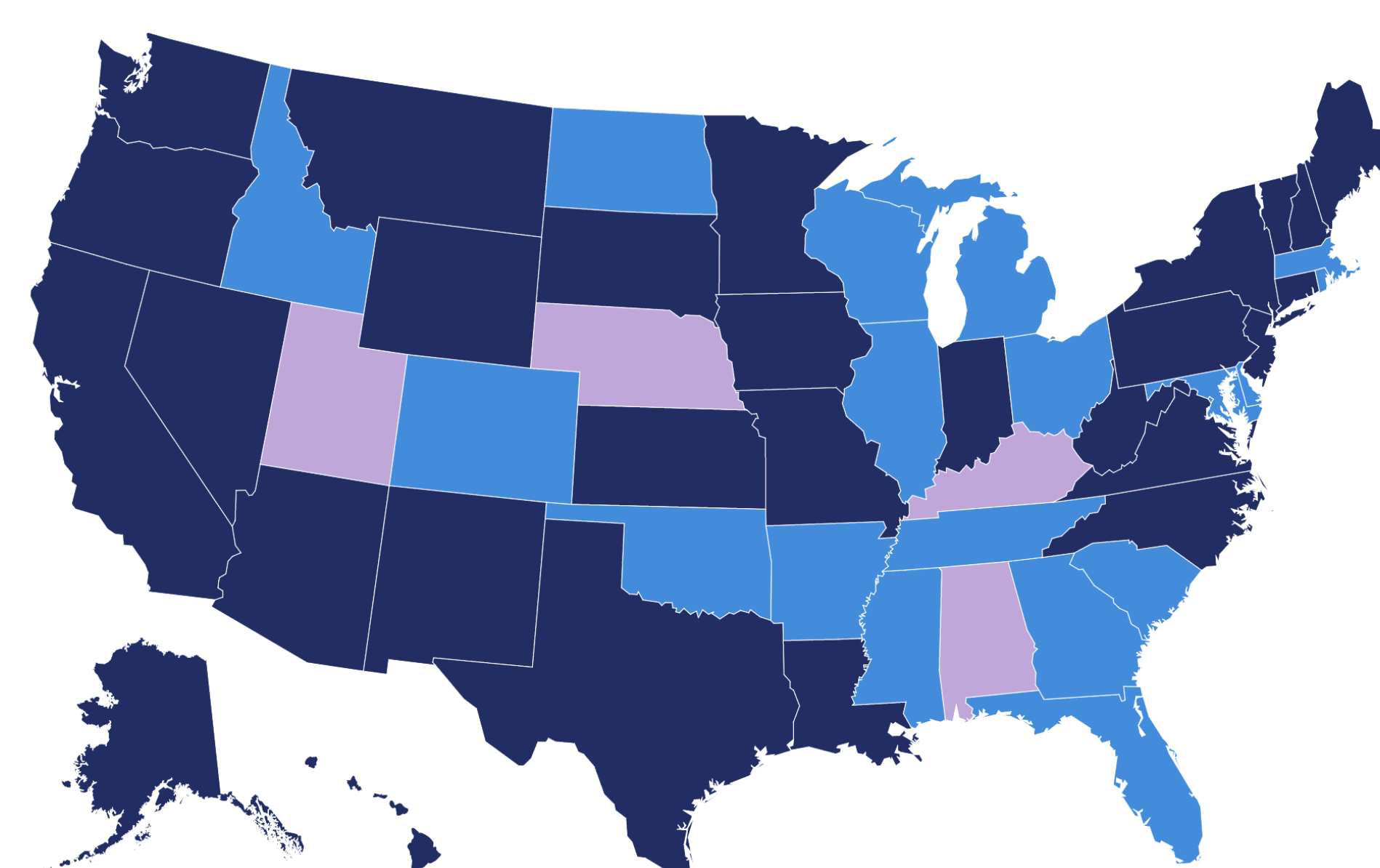
Fee-For-Service Data



Comprehensive Managed Care Data



Co



20

Person-level Exclusions Made, 2014-2020

Only beneficiaries in jurisdiction-plans included in the year are considered. There is overlap among these groups.

