

Investigating the Availability of Substance Use Disorder (SUD) Data in State APCDs

Statewide all-payer claims databases (APCDs) collect detailed information from health plans – including commercial, Medicaid, and Medicare payers – about their members and the specific services provided to patients. These valuable resources help state agencies, approved analysts, and others explore how healthcare use, prices, and outcomes vary over time and across a wide range of demographic, geographic, and other factors.

This includes examining the prevalence and treatment of patients with mental health (MH) and substance use disorder (SUD) conditions. Onpoint's most recent work on this front is featured in the dashboards developed for the Multi-State Behavioral Health Initiative, our innovative collaboration with five state APCDs.

For many years, researchers and policymakers often have been reluctant to pursue SUD-related studies using APCD data due to the widespread impression that these data are redacted, highly limited, or simply unavailable. This perception stems, in part, from the fact that many payers have different interpretations of federal privacy regulations regarding SUD-related records, limiting the data that they report to APCDs. However, recent [updates to 42 CFR Part 2](#) – a federal law governing the use and disclosure of SUD-related records – have sparked renewed interest in assessing the availability of SUD-related data in APCDs.

In this issue brief, we outline the empirical method that we used to determine the availability of SUD-related data in our client's

APCDs and how we evaluated whether the removal of SUD-related claims impacted the prevalence and utilization estimates for patients with mental health conditions, which often co-occur with SUD.

RECOMMENDATIONS FOR RESEARCHERS

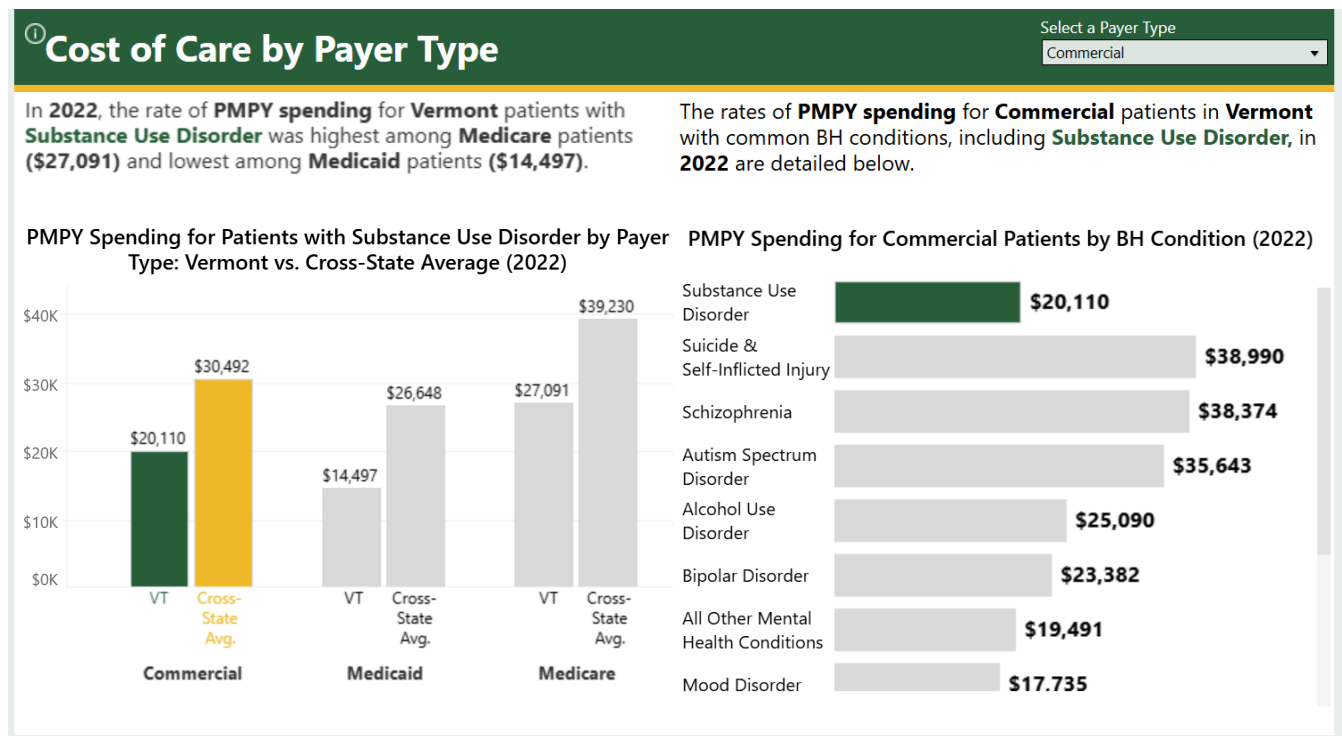
1. **The APCD may have more SUD-related data than expected.** While some payers do indeed restrict their reporting of SUD-related data, our analysis found that some APCDs have nearly 100% of their expected volume of SUD-related claims.
2. **Evaluate SUD-related data availability prior to beginning work.** Using the methodology described in this issue brief, analysts can estimate the general proportion of an APCD's SUD-related data that may be missing.
3. **Focus on payers with the most robust data.** Depending on study design, it may be possible to use only those payers with more complete SUD-related data and extrapolate findings for a larger population through imputation.
4. **Account for the connection between SUD and mental health diagnoses.** Since these conditions frequently co-occur, the wholesale removal of SUD-related claims may impact the reporting of other diagnoses and utilization for patients with both MH and SUD diagnoses.

DETERMINING THE FEASIBILITY OF SUD-RELATED STUDIES USING APCD DATA

Understanding how payers are reporting their SUD-related data to our clients' APCDs is a key part of Onpoint's comprehensive and ongoing data quality evaluations. These analyses have revealed both significant variation in payers' interpretation of 42 CFR Part 2 as well as differences in how payers that operate in different states vary their submission approaches according to state-specific regulations and requirements. For example, certain states require personal identifiers to be hashed prior to submission – a form of de-identification in which sensitive data are transformed into an indecipherable string of random characters. While some payers regard this requirement as sufficiently protective of patient privacy and

submit their SUD-related claims, others continue to withhold their SUD-related data.

Onpoint leveraged these ongoing assessments of SUD data availability to inform our initial steps when building the data set for the Multi-State Behavioral Health Initiative dashboards (see the figure below for an example from Vermont). Analyses of the SUD-related data in the APCDs of our five partnering states – Connecticut, Maryland, Minnesota, Vermont, and Washington – revealed that while some of the participating APCDs provide a full or nearly complete view of SUD-related claims, others offer a more limited data set.



To identify patients with SUD-related diagnoses, Onpoint's analysts utilized ICD-10-CM diagnosis codes reported in the medical claims. (International Classification of Diseases (ICD) diagnosis codes, which are maintained by the World Health Organization, serve as an international standard for the identification, recording, and reporting of highly specific diagnoses and medical conditions and are captured in the claims data submitted to APCDs.)

Onpoint specifically drew upon ICD code lists from the U.S. Substance Abuse and Mental Health Services Administration (SAMHSA) that provide a consistent definition of diagnoses related to SUD. (For additional detail, see Appendix A in SAMHSA's "[Mental Health Annual Report 2015-2020 - Use of Mental Health Services: National Client-Level Data](#).”) We then calculated the proportion of medical claims that had a primary diagnosis related to SUD by state, individual

payer, and payer type (i.e., commercial, Medicaid, Medicare).

Next, Onpoint focused on the data submitted by payers known to remove some or all of their SUD-related information, comparing the percentage of these payers' SUD-related medical claims to the percentage from payers that had reported no similar restrictions. The results were used to establish a benchmark range for the percentage of a payer's total medical claims that could be expected to be related to SUD if the reporting of those claims were not restricted.

For each state, we were able to compare data from each payer against the benchmark to identify those payers that were removing some or all SUD-related claims from their APCD submissions. Once payers who were removing SUD claims were identified, an analysis of these payers' market share both within the state and within their respective payer type (commercial, Medicaid, Medicare) was used to estimate the likely volume of missing SUD-related data for each state and payer type.

DETERMINING THE FEASIBILITY OF SUD-RELATED STUDIES USING APCD DATA

For **commercial** payers that reported submitting all of their SUD-related data, the minimum percentage of claims with a primary diagnosis related to SUD was 0.3% (with many payers showing significantly higher rates — some as high as 2%). For **Medicaid**, our assessment established a minimum threshold of 2.5% (with certain payers reporting more than 8% of their claims as related to SUD). For **Medicare**, the minimum threshold was lower – 0.15% – though some payers reported nearly 2% of their claims as having an SUD-related primary diagnosis. As expected, payers that indicated their withholding of SUD-related claims reported rates below these established thresholds.

Based on this analysis, we classified individual payers into three categories: (1) SUD-related data is available, (2) SUD-related data may contain gaps, and (3) SUD-related data is not available (table below).

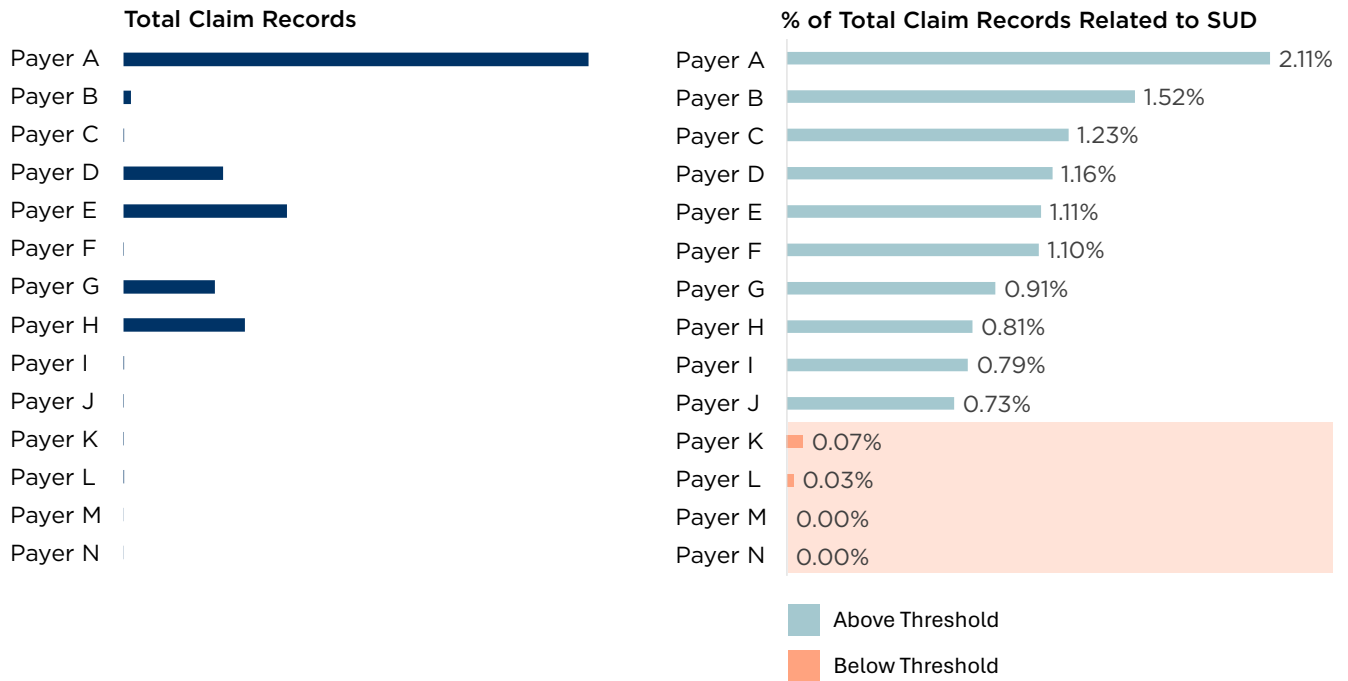
Benchmarks for Determining SUD-Related Data Availability

Data Availability Category	Percentage of SUD-Related Claims by Payer Type)		
	Commercial	Medicaid	Medicare
SUD-related data is available	≥0.3%	≥2.5%	≥0.15%
SUD-related data may contain gaps	0.1% < - <0.3%	1.0% < - <2.5%	0.1% < - <0.15%
SUD-related data is not available	≤0.1%	≤1.0%	≤0.1%

Once we applied these benchmarks at the payer level and categorized each payer based on their SUD-related data availability, we analyzed each category in combination with the total number of claims submitted by each payer to approximate the percentage of SUD-related data likely to be missing.

The following graph offers an example of the volume and variation in SUD-related data completeness among commercial payers within a single APCD. (Please note that state names are not included in this issue brief.)

Example of Total Medical Claim Records Submitted by Payers Compared to the Percentage of SUD-Related Records (Based on Primary Diagnosis & SAMHSA Categories, Counts Blinded) (2022)



The following table offers an illustration of how Onpoint estimated the percentage of SUD-related data available by payer type. This estimation incorporated both the share of records submitted by each payer and an evaluation of whether the payer was including or excluding SUD-related claims.

Commercial Payer Example of Calculation of Percentage of Available SUD Data

Commercial Payer	% of Claims Related to SUD	Above the Commercial Threshold (0.3%)?	Total Records from This Payer	Sum of Records from Payers that Included SUD Data
A	2.5%	Yes	1,000	1,000
B	1.7%	Yes	20,000	20,000
C	0.5%	Yes	7,000	7,000
D	0.04%	No	12,000	--
			40,000	28,000

Next, to determine a state or payer type’s SUD-related data availability, we determined the percentage of the claims in a state’s APCD that came from payers that were not removing SUD-related data. Continuing with the example detailed in the table above, we would use the following calculation:

$$\begin{aligned}
 &\text{Estimated \% of SUD-related data availability} = \\
 &\frac{\# \text{ of records (SUD and non-SUD) from plans meeting threshold}}{\text{total \# of records (SUD and non-SUD) from all plans}} = \\
 &\frac{28,000}{40,000} = \\
 &\mathbf{70\%}
 \end{aligned}$$

IDENTIFYING DIFFERENT PAYERS' APPROACHES TO SUD-RELATED CLAIMS REMOVAL

Further investigation into SUD-related categories and diagnoses revealed variations in how payers were selectively withholding specific types of SUD-related data. For instance, one payer might include claims with alcohol-related diagnoses in their submissions (Payer A in the illustration below) while removing claims that had diagnoses that fell into other SUD-related categories such as nicotine or opioids. Another payer (Payer B), however, might include nicotine-related claims but exclude all other SUD-related categories. This inconsistent approach to data submission helps explain some of the variation observed in payers that fell below the established thresholds and adds further complexity to SUD-related analyses.

Difference in Payers' Approach to SUD-Related Claims Removal (Example)

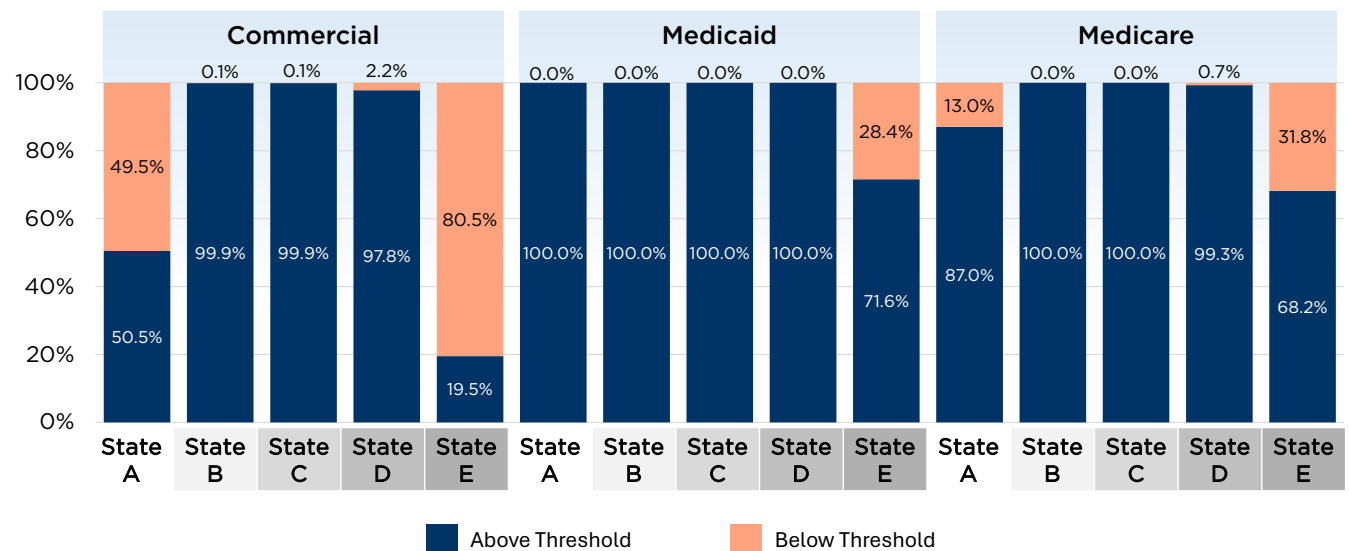


RESULTS

We applied our benchmarks to payers in each of the five states participating in the behavioral health study to estimate the percentage of APCD claims that were provided by payers who removed SUD data versus those who did not remove SUD-related data. For example, State A had 50.5% of commercial data from payers who

were above the threshold and were likely submitting their SUD data, while 49.5% of its commercial data was from payers who were below the threshold and were likely their SUD data. Using these methods, Onpoint's health data analysts estimated the percentage of SUD-related data completeness by payer type and state (see figure below).

Estimated % of Medical Claims in the APCD from Payers Above & Below the SUD-Related Threshold (2022)



Among the five participating states, we estimated that three APCDs had complete or nearly complete SUD-related data across the reporting years. Of the two remaining states, one showed significantly lower rates of SUD-related data completeness, with only 20% of the commercial market estimated to include all of their SUD-related claims. The other's Medicaid data appeared complete while some removal of SUD-related data was observed among its commercial and Medicare payers. In these two states, there appeared to be substantial removal of SUD-related data by payers representing a significant portion of total claims volume.

Based on these findings, we concluded that our study could include SUD-related reporting and cross-state benchmarking for the three states that had complete or nearly complete SUD-related data. Due to the less complete SUD-related data in the other two states, we limited their state-specific reporting to mental health analyses only and excluded SUD-related analysis.

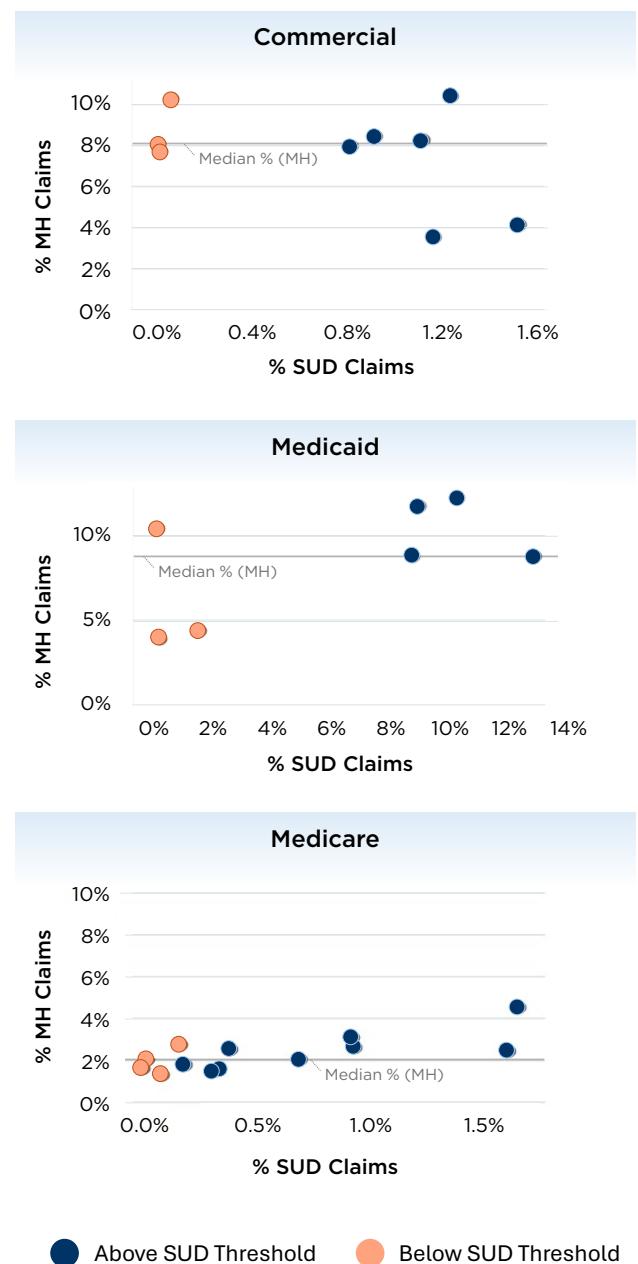
EXPLORING THE RELATIONSHIP BETWEEN SUD & MENTAL HEALTH DISORDERS IN CLAIMS

A key consideration when studying mental health (MH) in states where APCDs lack adequate SUD-related data is the potential for SUD-related claim removal that results in lower observed utilization rates of MH-related services. This is particularly relevant for patients with co-occurring MH- and SUD-related conditions. As part of our survey work to understand how different payers handle the exclusion of their SUD-related data, several interview respondents indicated that SUD-related claims often are excluded entirely, rather than simply removing only the SUD-related information (e.g., the diagnosis code). If this approach is widespread, we would expect states with less comprehensive SUD-related data to also show reduced rates of MH-related diagnoses and care utilization for patients with co-occurring conditions.

To conduct a sensitivity analysis on the potential impact of SUD-related claim removal on the rates of MH disorder prevalence and treatment, we focused on commercial, Medicaid (both traditional and managed care), and Medicare (both traditional and managed care) payers in one of the participating states. We analyzed the percentage of MH-related claims by payer and insurance type, comparing plans below the SUD-related claims threshold (SUD-related claims likely removed) to those above the threshold (SUD-related claims likely included).

The results of this analysis are shown in the graphs to the right.

Percentage of MH Claims Above & Below the SUD-Related Threshold by Payer Type & Payer (2022)



At this aggregate level, we observed no significant association between the removal of SUD-related claims and the percentage of claims with a primary MH diagnosis. Plans removing their SUD-related claims nevertheless reported percentages of MH diagnoses that were similar to those of plans that did not remove SUD-related claims. This suggests that while the removal of SUD claims may marginally decrease the percentage of reported claims related to mental health, the overall impact may not be substantial.

Additionally, Onpoint conducted an analysis of co-occurring MH- and SUD-related diagnoses on claims. For payers that submitted SUD-related claims to the APCD, we also analyzed additional diagnoses listed on the claims beyond the primary (or leading) diagnosis to determine the percentage of claims that had both SUD- and MH-related diagnoses. Between 4.0% and 5.5% of all behavioral health claims submitted by commercial payers had co-occurring MH and SUD diagnoses. The ranges were higher for behavioral health claims for both Medicaid (6.6% - 9.0%) and Medicare (4.6% - 12.7%).

RECOMMENDATIONS FOR ANALYSTS & RESEARCHERS

- 1. The APCD may have more SUD data than expected.** Many researchers may assume that state APCDs do not receive SUD data, rendering them unsuitable for SUD-related studies. However, our findings suggest otherwise. In fact, some state APCDs appear to include nearly 100% of the anticipated SUD data from claims, highlighting their potential for SUD-related research.
- 2. Evaluate SUD data availability prior to initiating your study.** The state agency overseeing an APCD likely can offer key insights into the completeness of its SUD-related data, including guidance regarding the extent of available data. If such information is unavailable, analysts can conduct their own evaluation of SUD-related data availability. If taking this approach, our recommendations include:
 - Calculating the percentage of claims with an SUD-related primary diagnosis by payer type (i.e., commercial, Medicaid, Medicare) and by individual payer to identify those above or below the expected threshold and completion range
 - Examining whether specific SUD-related diagnoses (e.g., alcohol, cannabis, nicotine) are being included or restricted from reporting by particular payers
 - Determining the total claims volume submitted by payers that are excluding SUD-related data to estimate the proportion of missing SUD-related data across payers; carriers with a larger volume of covered lives will have a more significant impact
- 3. Focus on payers with the most robust data and consider imputation for those that restrict their SUD-related information.** If certain payers remove SUD-related data while others provide it, analysts can consider excluding payers that are removing relevant data. Depending on study design, it may be possible to use a subset of the population or extrapolate findings for the full population through imputation. This approach may be preferable to either abandoning the study entirely or utilizing data from payers who remove a significant volume of their SUD-related information.
- 4. Account for the removal of SUD-related claims in mental health findings.** A small percentage of claims include both SUD and MH diagnoses. Based on feedback from payers, those excluding SUD-related data may be removing entire claims instead of just the data points related to SUD. Since a subset of patients have co-occurring MH and SUD conditions, such removals could cause slightly lower observed rates of MH-related diagnoses and service utilization

in APCD-based analyses. It is important to note these potential limitations when interpreting mental health findings if SUD-related data is removed. However, as the proportion of claims with diagnoses related to both mental health and SUD is relatively small, this should not deter researchers from conducting MH-focused studies using APCD data.

CONCLUSION

The prevalence and treatment of substance use disorders are of increasing interest to both public health officials and policymakers as important areas of analysis. Onpoint's recent evaluation shows that, contrary to a common misconception among researchers, state APCDs are a valuable resource for those wishing to focus on substance use disorder in communities and the healthcare system. However, given the variation in how different payers interpret relevant federal and state privacy restrictions, it remains important for researchers to evaluate the availability of SUD-related data prior to starting their analytic projects.

ABOUT THE AUTHOR



For questions or further information, please contact Onpoint's Director of Health Analytics, Amy Kinner, MS, at akinner@onpointhealthdata.org.

Additional contributions were provided by Onpoint's Senior Data Analytics Manager, Gina Robertson.

ABOUT ONPOINT HEALTH DATA

Onpoint Health Data is a nonprofit organization that specializes in collecting, integrating, and analyzing health data to provide our clients with enriched data sets and innovative analytic solutions tailored to their specific needs. We are an independent, nonpartisan organization supporting federal, state, and regional health improvement initiatives for nearly 50 years.



Learn
More



Contact
Onpoint



Join Us
on LinkedIn