ASPE
ASSISTANT SECRETARY FOR PLANNING AND EVALUATION

OFFICE OF HUMAN SERVICES POLICY

BRIEF
November 2023

Advancing Research on Intersections of Child Welfare and Medicaid Using Linked Data from the CCOULD Project

Brett Greenfield, Emily Madden, and Robin Ghertner

Increasing availability of linked child welfare and Medicaid data can advance research on the intersections of child welfare and Medicaid. The project, *Child and Caregiver Outcomes Using Linked Data* (CCOULD), developed a research-use dataset combining child welfare records and Medicaid claims for children and families involved in child welfare systems in Florida and Kentucky. This brief aims to increase usage of CCOULD data by researchers, drawing attention to the capacity of the data to answer pertinent research questions, and assist researchers in maximizing the utility of the data and navigating limitations.

KEY POINTS

- Research at the nexus of child welfare system engagement and Medicaid utilization has many benefits including, improving services coordination for children and families engaged in these systems, preventing child abuse and neglect and child welfare system involvement, and improving health and child well-being outcomes.
- The CCOULD dataset consists of 16 files containing specific information pertaining to child welfare engagement or Medicaid utilization from Kentucky and Florida, over the 2016-2021 period. The data include records for over 1,000,000 children, and over 85,000 caregivers with children engaged with child welfare services.
- Among others, research topics that CCOULD can be used to address include:
    - The prevalence of substance use disorders (SUD) among children and caregivers engaged in child welfare systems.
    - The relationship between SUD or mental health conditions and child welfare system engagement.
    - The relationship between services provided through Title IV-E (e.g., family support, case management, SUD treatment, etc.) and child and caregiver child welfare and health outcomes.
    - Medicaid spending for child welfare engaged families.

INTRODUCTION

A large proportion of children who receive child welfare services are eligible for Medicaid. Linking Medicaid data to child welfare system data can identify opportunities to improve coordination between these two programs and improve the effectiveness and cost-effectiveness of their services, among other objectives. The *Child and Caregiver Outcomes Using Linked Data* (CCOULD) project provided technical assistance to two states – Kentucky and Florida – in linking their Medicaid administrative claims of children and their caregivers with case-level data from the child welfare system. CCOULD then combined state-specific datasets by mapping them onto a common data model and created a multi-state, deidentified dataset for secondary data analysis by researchers.

CCOULD generated a number of important lessons for how states can link data across child welfare and Medicaid systems. These lessons, as well as other information on the data, are documented in a separate ASPE report, "Linking Child Welfare and Medicaid Data: Lessons Learned from Two States".[1] Details on the variables, structure, data quality, and other documentation are available through the National Data Archive on Child Abuse and Neglect (NDACAN).[2]

The CCOULD project and data demonstrates the importance of research at the nexus of child welfare engagement and Medicaid usage. Linked data on services from child welfare and Medicaid are valuable for understanding a range of questions and topics that address critical challenges faced by families at risk of child welfare system involvement. Such research includes identifying indicators of risk for child abuse and neglect to better target prevention efforts, improving the implementation and effectiveness of interventions designed to treat substance use disorders and other behavioral health challenges, as well as promoting child and family well-being.

This brief discusses a range of potential research topics that could be explored using the CCOULD data, with the goal of informing and inspiring researchers to engage critically with this innovative and robust dataset to improve the lives of children and families. Additionally, this brief provides guidance to researchers regarding both the capacity of the data to address essential research questions, as well as limitations of the data to aid researchers in their use of these data.

OVERVIEW OF DATA

The CCOULD dataset consists of 16 data files each containing specific information pertaining to child welfare engagement or Medicaid utilization. Children engaged with child welfare systems in Kentucky and Florida from 2016 to 2021, and their caregivers were included in the dataset. Data files can be merged, and children can be linked to caregivers when present in multiple data files. Instructions for merging data files can be found in the NDACAN documentation. Below are highlights of the available data:

- Information on child welfare system engagement (including allegations and substantiated allegations of child maltreatment) on over 1,000,000 children.
- Information on child welfare services supported in part by federal funds through Title IV-E of the Social Security Act. Note, Title IV-E prevention services data are not yet uniformly collected by states, and some of these data were collected prior to any collection of Title IV-E prevention services data. Certain services may be undercounted.
- Information on foster care episodes for over 95,000 children.
- Information on Medicaid enrollment, diagnoses, inpatient and outpatient claims, and prescription data on over 850,000 children with any child welfare involvement included in

---

[1] https://aspe.hhs.gov/reports/ccould-lessons-learned-report

[2] https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/dataset272codebook.pdf

the child welfare report data (including allegations and substantiated allegations of child maltreatment), and over 85,000 caregivers with children engaged with child welfare services.

- All Medicaid files have a corresponding randomly selected comparison sample of non-child welfare engaged individuals that can be used for analyses.
- Demographic data (race, ethnicity, and sex)[3] are available for linked children and caregivers, as well as comparison samples. Note, sexual orientation data are not included.

RESEARCH TOPICS USING CCOULD

This section outlines potential research topic areas and research questions that could be examined using the dataset. These research questions are intended to help researchers get oriented to the data, and do not reflect the full range of questions that could be addressed. Once working with the actual data, researchers will need to refine their specific research questions based on the specific strengths and limitations of the variables and subpopulations of interest. The table below describes topical areas for research, brief insight into the types of variables available in the data that may be useful, as well as potential limitations.

Research Topics	CCOULD variables and source: 1 = Medicaid claims 2 = Child Welfare data
Prevalence of substance use disorders among children and caregivers engaged in child welfare systems How do claims-based indicators of substance use disorders (SUD) in Medicaid data (i.e., diagnosis codes) compare with indicators of SUD in the child welfare data? Is one system more likely to identify SUD, and make connections to services? Among people covered by Medicaid, how does the prevalence of SUD compare across children and caregivers based on child welfare system engagement? Are certain types of SUD (e.g., opioid, alcohol) more commonly diagnosed among Medicaid recipients with child welfare involvement compared to those without? What percentage of children in the foster care system have a parent with a SUD diagnosis on Medicaid claims and/or child welfare (CW) data?	· SUD diagnosis codes (1) · SUD-related services provided by CW (under Title IV-E) (2) · Healthcare Common Procedure Coding (HCPC) and revenue codes to identify SUD treatment (1) · National Drug Codes (NDCs) for medications to treat SUD (1) · ID variable to link parents/caregivers in Medicaid to children in the CW data (1, 2)· Caseworker determinations of the role of alcohol and drugs (2)

---

[3] Demographic variables were constructed for the research use dataset. "Race" includes: American Indian or Alaskan Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, and White. "Ethnicity" includes: Hispanic or Latino (yes/no). "Sex" includes: male and female, and is labeled "gender" in the dataset.

How frequently are children with identified fetal alcohol spectrum disorders (FASD) in the child welfare system? How do rates of FASD diagnoses compare among non-child welfare engaged children? How do child welfare outcomes differ for children with identified FSAD compared to those without? What are the demographic characteristics of parents with alcohol use disorder (AUD) in the sample? How does access to treatment for parents with AUD differ for those with children engaged in child welfare systems compared to those without? How do prevalence rates of SUDs among child welfare engaged children and caregivers differ by demographic characteristics? **Considerations in the data:** • Reliance on accuracy of International Classification of Disease (ICD) codes • Requires cross-referencing of ICD, Healthcare Common Procedure Coding (HCPC) & National Drug Codes (NDCs) • < 3% of cases include Title IV-E SUD-related services • Child welfare reports include data on caregiver substance use in < 25% of cases (~400,000 records), and does not require a SUD diagnosis • Requires use of all Medicaid data files to identify possible diagnoses • Race and ethnicity information relies on child welfare records, and there is high percentage of missing data for child ethnicity	· Demographic characteristics of parents/caregivers (e.g., age, race, disability status, rural/urban) [1, 2] · Length of time in foster care, maltreatment incidences, and measures of re-entry (2)
Relationship between SUD or mental health conditions and child welfare engagement and outcomes How do child welfare outcomes for children with parents with SUDs (e.g., out-of-home removal rates, reunification rates, length of time in out-of-home care) compare to those without identified SUDs? How do these outcomes differ by child demographic characteristics? How do child welfare outcomes differ among children whose parents have different types of SUDs (e.g., Opioid use disorder, alcohol use disorder, etc.)? How frequently are parents and caregivers receiving SUD treatment involved in the child welfare and/or Medicaid systems?	· HCPC and revenue codes to identify SUD treatment (1) · SUD treatment codes (2) · NDCs for medications to treat SUD (1) · Clinical characteristics (e.g., mental health diagnosis, hospitalization or ED visit for SUD or overdose) [1] · ID variable to link parents/caregivers in

What type of SUD treatment are parents and caregivers receiving? To what extent are postpartum depression or other mental health conditions associated with child welfare involvement, care experiences, and outcomes? How do child welfare outcomes differ among children of caregivers with mental health diagnoses compared to those who do not? How do child welfare outcomes differ for children whose caregivers are receiving mental health treatment compared to children of caregivers who are not? How does receipt of mental health treatment for mental health conditions differ by caregiver demographics including age, race, and gender? How does psychotropic medication usage differ among child welfare engaged children and adults compared to those without child welfare engagement? **Considerations in the data:** • Reliance on accuracy of ICD codes • Requires cross-referencing of ICD, Healthcare Common Procedure Coding (HCPC) & National Drug Codes (NDCs) • Certain diagnostic codes (e.g., Fetal Alcohol Syndrome) appear more commonly in secondary or tertiary diagnoses, and are identified alongside other primary diagnoses • < 3% of cases include Title IV-E SUD-related services • Requires use of all Medicaid data files to identify possible diagnoses • Child welfare reports include data on caregiver substance use in < 25% of cases (~400,000 records) • Child placement data available for ~95,000 children • Race and ethnicity information relies on child welfare records, and there is high percentage of missing data for child ethnicity	Medicaid to children in the CW data (1, 2) · SUD diagnosis codes (1) · Length of time in foster care, maltreatment incidences, and measures of re-entry (2) · Mental health-related services provided by CW (under Title IV-E) (2)
Composition of substance use disorder treatment recipients What demographic characteristics, clinical characteristics, and medical care utilization patterns are associated with parents/caregivers receiving Medicaid-funded SUD treatment? What demographic characteristics are associated with parents/caregivers receiving Title IV-E funded SUD treatment?	· Demographic characteristics of parents/caregivers (e.g., age, race, disability status, rural/urban) [1, 2] · Clinical characteristics (e.g., mental health diagnosis, hospitalization

Considerations in the data: • Reliance on accuracy of ICD codes • Requires cross-referencing of ICD, Healthcare Common Procedure Coding (HCPC) & National Drug Codes (NDCs) • Requires use of all Medicaid data files to identify possible diagnoses, and use of services & medications • < 3% of cases include Title IV-E SUD-related services • Demographic data relies on child welfare records for matched children and caregivers, and there is high percentage of missing data for child ethnicity • Demographic data for Medicaid comparison datasets is collected less systematically	or ED visit for SUD or overdose) [1] · Medical care utilization patterns (e.g., receipt of primary care) [1] · HCPC and revenue codes to identify SUD treatment (1) · SUD-related services provided by CW (under Title IV-E) (2) · NDCs for medications to treat SUD (1)
Impact of SUD treatment on child welfare outcomes How does receipt of SUD treatment by parents/caregivers with children involved with child welfare systems relate to child welfare outcomes, such as length of time in foster care, repeat maltreatment incidences, and re-entry into foster care? How does having a parent receiving SUD treatment affect out-of-home removal and type of placement decisions? How does receipt of SUD treatment by parents/caregivers with children engaged in child welfare systems differ by caregiver demographic characteristics? **Considerations in the data:** • Reliance on accuracy of ICD codes • Requires cross-referencing of ICD, Healthcare Common Procedure Coding (HCPC) & National Drug Codes (NDCs) • < 3% of cases include Title IV-E SUD-related services • Requires use of all Medicaid data files to identify possible diagnoses • Requires accounting for multiple foster care incidences per child • Child welfare reports include data on caregiver substance use in < 25% of cases (~400,000 records) • Child placement data available for ~95,000 children • Demographic data relies on child welfare records for matched children and caregivers, and there is high percentage of missing data for child ethnicity	· HCPC and revenue codes to identify SUD treatment (1) · SUD-related services provided by CW (under Title IV-E) (2) · NDCs for medications to treat SUD (1) · Length of time in foster care, maltreatment incidences, and measures of re-entry (2) · Demographic characteristics of parents/caregivers (e.g., age, race, disability status, rural/urban) [1, 2]

Relationship between services provided through Title IV-E prevention services (e.g., mental health, substance abuse, and in-home parent skills) and child and caregiver child welfare and health outcomes Which Title IV-E prevention services are used most often among child welfare involved families? How does use of Title IV-E prevention services differ among caregivers of various demographic groups? What is the relationship between receipt of Title IV-E prevention services and caregiver health outcomes (e.g., increase in primary care visits, indicators of improved health, etc.) **Considerations in the data:** • Requires use of all Medicaid files for cross-referencing of Healthcare Common Procedure Coding (HCPC) & National Drug Codes (NDCs) • Low percentage of cases documented as receiving certain Title IV-E prevention services due to timing of data collection • Other child welfare services not funded by Title IV-E (e.g., housing, employment) may also be included in the data, but are not uniformly collected by all states • Requires accounting for multiple foster care incidences per child • Child placement data available for ~100,000 children • Demographic data relies on child welfare records for matched children and caregivers, and there is high percentage of missing data for child ethnicity	· HCPC and revenue codes to identify SUD treatment (1) · NDCs for medications to treat SUD (1) · Title IV-E SUD treatment codes (2) · Length of time in foster care, maltreatment incidences, and measures of re-entry (2) · Demographic characteristics of parents/caregivers (e.g., age, race, disability status, rural/urban) [1, 2]
Use of services for children with substance use related illnesses What services funded by Medicaid and child welfare systems are used by children born with substance use related illnesses, and their parents? How are those services related to child health and child welfare outcomes? What differential diagnoses in children are most commonly associated with parental SUD? How do substance use related illnesses differ among children of different demographic groups, whether engaged in child welfare systems or not? **Considerations in the data:**	· Diagnosis codes (1) · Length of time in foster care, maltreatment incidences, and measures of re-entry (2) · SUD-related services provided by CW (under Title IV-E) (2) · Length of time in foster care, maltreatment incidences, and measures of re-entry (2)

• Reliance on accuracy of ICD codes • Requires cross-referencing of ICD, Healthcare Common Procedure Coding (HCPC) & National Drug Codes (NDCs) • < 3% of cases include Title IV-E SUD-related services • Requires accounting for multiple foster care incidences per child • Child placement data available for ~95,000 children • Demographic data relies on child welfare records for matched children and caregivers, and there is high percentage of missing data for child ethnicity • Demographic data for Medicaid comparison datasets is collected less systematically	· Demographic characteristics of parents/caregivers (e.g., age, race, disability status, rural/urban) [1, 2]
Medicaid spending for child welfare engaged families How much does Medicaid spend on SUD treatment among parents with children in the foster care system? How much does Title IV-E spend? Does spending on parent SUD treatment result in subsequent savings to Medicaid and/or Title IV-E? Does spending on SUD treatment and/or Title IV-E spending affect foster care spending? How long do children retain Medicaid enrollment following exits from foster care? **Considerations in the data:** • Reliance on accuracy expenditure & enrollment data • Longitudinal analyses restricted 5 years of data included	· SUD expenditure data on Medicaid claims (1) · Medicaid enrollment data (1) · SUD expenditure data on CW invoices (2) · Longitudinal data to tracking individual parents and their children over time (2) · Variable to link of parents to children (1, 2)
Health and mental health outcomes for child welfare engaged children What diagnoses are most common among children engaged in child welfare systems? How do health outcomes for children engaged in child welfare systems differ by demographic characteristics? What diagnoses are associated with different forms of maltreatment? How are the timing of diagnosis and treatment of health problems in children related to timing of child welfare engagement? How do certain health problems affect child welfare outcomes?	· Diagnosis codes (1) · Medical care utilization patterns (e.g., receipt of primary care) [1] · Medicaid enrollment data (1) · NDCs, HCPC and revenue (1) · Length of time in foster care, maltreatment incidences, and measures of re-entry (2)

What health diagnoses and healthcare services are received by children in child welfare congregate care settings such as group homes and residential treatment settings? Does Medicaid data provide insights as to the prevalence or severity of the health issues of children in these care settings? **Considerations in the data:** • Reliance on accuracy of ICD codes • Requires cross-referencing of ICD, Healthcare Common Procedure Coding (HCPC) & National Drug Codes (NDCs) • Requires use of primary, secondary, and tertiary diagnoses to accurately determine scope of care • Requires accounting for multiple foster care incidences per child • Child placement data available for ~95,000 children • Demographic data relies on child welfare records for matched children and caregivers, and there is high percentage of missing data for child ethnicity	· Demographic characteristics of parents/caregivers (e.g., age, race, disability status, rural/urban) [1, 2]

NEXT STEPS FOR RESEARCHERS

This new and innovative dataset is poised to advance knowledge about children and families engaged with child welfare and Medicaid systems. This brief aims to spur researchers to begin using these data. The scope of research questions that can be addressed with these data extend beyond what is described in this brief. ASPE is committed to supporting efforts to link data systems to improve the lives of children and families, and aims for this project and the research stemming from it to inspire states and jurisdictions to build similar data infrastructures that align with these goals. For more information about the work ASPE is doing to further assist states with linking child welfare and Medicaid data, please read more about the Child Welfare and Health Infrastructure for Linking and Data Analysis of Resources, Effectiveness, and Needs (CHILDREN) Initiative.

Researchers interested in conducting research using data developed from the CCOULD project can request access to the data directly from NDACAN. Access to the data is free of charge, but requires compliance with NDACAN's standards for data security and approval from an Institutional Review Board (IRB). ASPE plans to support researchers conducting important research at the intersection of child welfare and Medicaid using the CCOULD dataset through a research consortium. More information about the call for proposals under the research consortium can be found here.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Office of the Assistant Secretary for Planning and Evaluation

200 Independence Avenue SW, Mailstop 447D
Washington, DC 20201

For more ASPE briefs and other publications, visit:
aspe.hhs.gov/reports

ABOUT THE AUTHORS

Brett Greenfield is a Society for Research in Child Development/AAAS Federal Policy Fellow in the Data and Technical Analysis Division in the Office of Human Services Policy in the Office of the Assistant Secretary for Planning and Evaluation.

Emily Madden is a Social Science Research Analyst in the Data and Technical Analysis Division in the Office of Human Services Policy and Office of the Assistant Secretary for Planning and Evaluation.

Robin Ghertner is Director of the Data and Technical Analysis Division in the Office of Human Services Policy and Office of the Assistant Secretary for Planning and Evaluation.

ACKNOWLEDGEMENTS

The RTI International project team was led by Tami L. Mark, Melissa Dolan, Benjamin Allaire, & Christina Bradley. Emily Madden (ASPE) was the Contracting Officer’s Representative, and Valeria Butler (OPRE) was the project lead.

SUGGESTED CITATION

Greenfield, B., Madden, E., & Ghertner, R. (2023). Advancing Research on Intersections of Child Welfare and Medicaid Using Linked Data from the CCOULD Project. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services.