

Medicaid Data Linkages for Health Outcomes Research

The Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) supports U.S. Department of Health and Human Services (HHS) projects that focus on **improving data capacity – the collection, linking, and analysis of data** – for health outcomes research. Through the OS-PCORTF, ASPE has funded 11 projects that sought to link Medicaid data with other data sources to create **more comprehensive, longitudinal datasets** that overcome the limitations of using Medicaid data in isolation. This resource provides a snapshot of those projects, and key considerations for conducting future Medicaid linkages.

Projects produced 6 linked datasets.

Five are available to external researchers, while one* is limited to state use only.

- [Child and Caregiver Outcomes Using Linked Data \(CCOULD\)](#)
- [Linked National Hospital Care Survey \(NHCS\) to Medicaid Data](#)
- [Medicaid Enrollee Supplemental File: National Death Index Segment](#)
- [Multi-State Network of Linked Pregnancy Risk Assessment Monitoring System \(PRAMS\) and Clinical Outcomes Data*](#)
- [National Ambulatory Medical Care Survey \(NAMCS\) Health Center Component to Centers for Medicare & Medicaid Services \(CMS\) Medicaid Data](#)
- [Privacy Preserving Record Linkage to CMS](#)

Projects linked T-MSIS or State Medicaid data to:

- Clinical/electronic health record systems data
- Human services data
- Social and environmental data
- Survey data
- Surveillance data
- Vital records (birth, mortality)

Data Linkage Identifiers

Projects used one or more of the following types of identifiers



Personally identifiable information (PII) that directly identifies an individual (ex: first and last name; Social Security number) (n=8 projects).



Other identifiers such as official numbers or codes assigned by entities like state agencies (ex: Medicaid ID) (n=8 projects).



De-identified tokens that are unique, encrypted identifiers created from PII to facilitate privacy-preserving record linkage (n=1 project).

Most projects used a **combination of deterministic and probabilistic methods** to link datasets.



Deterministic: Linkage that uses specific, predefined criteria to link records across different datasets. This method relies on exact matches of key identifiers.



Probabilistic: Linkage that uses statistical theory to estimate the likelihood that variables match across datasets based on the degree of agreement between variables.



Privacy-preserving record linkage (PPRL): Connects records across datasets while ensuring the privacy of individuals' information through encrypted coding techniques.

The OS-PCORTF linked Medicaid datasets have been used in **over 100 studies**. Linked Medicaid datasets have the potential to address previously understudied research domains and support policymaking.

100+
studies



Health Services and Utilization
Understanding when and how individuals interact with multiple systems



Maternal and Infant Health
Connecting mothers to their children to study continuum of care and life course trajectories



Methodological Advancements
Supporting more robust research designs and statistical approaches to develop more predictive, personalized, and efficient care solutions



Rare Diseases
Increasing population coverage and continuity to study rare conditions



Social and Behavioral Health
Examining social causes of diseases and offering a systems-level perspective on well-being



Policy Insights
Exploring the broader economic and policy impacts of Medicaid enrollment on health and other outcomes of beneficiaries

Considerations for Future Data Linkages

Strengthening data linkages at both the state and federal level is critical to advancing health services research that inform evidence-based policy development and program improvement.



Streamline processes for data sharing and acquisition

Projects often face time and legal barriers to securely accessing data for linkages. Encouraging data owners to develop clearer instructions and better pricing models can facilitate efficient data access.



Advance tools and techniques to securely exchange data

PPRL facilitates secure linkage of health information, but technical enhancements are needed to enable widespread implementation. Having a standardized tool for PPRL could create consistency and allow for more cross-agency linkages.



Document linkage processes to support replication

Project teams can maintain records and develop technical assistance resources, such as user guides, roadmaps, or common data models, to improve the efficiency of future linkage projects.



Provide datasets in multiple file formats to support various research needs

Some researchers require identifying variables in restricted-use files for their analyses, while others can use aggregate, deidentified data to answer research questions. Providing flexible data access options may increase the utilization of linked datasets.



Expand the utility of linked datasets to study a range of research use cases

Broadening the scope of existing linked datasets to include other clinical domains or populations can increase the usability and relevance of datasets for health outcomes research.



Engage key stakeholders, including patients, early and throughout the process

Project teams can involve policymakers, researchers, and subject matter experts through technical expert panels or stakeholder engagement sessions. Additionally, engaging the patient or caregiver population that the dataset will serve can boost its credibility and use.



Promote training and knowledge sharing to build capacity

States and federal agencies face staffing, funding, and infrastructure limitations. Training internal staff on linkage methodologies and encouraging knowledge sharing through consortia or communities of practice can help build capacity for data linkages.



For more information about the OS-PCORTF projects described in this resource, please reach out to OSPCORTF@hhs.gov

Explore the OS-PCORTF project portfolio at <https://aspe.hhs.gov/collaborationscommittees-advisory-groups/osp cortf/explore-portfolio>