

Sharing and Accessing Administrative Data: Promising Practices and Lessons Learned from the Child Maltreatment Incidence Data Linkages Project

Beth Varley and Claire Smither Wulsin

Introduction

Accurate and ongoing surveillance of the incidence of child maltreatment and related risk and protective factors can help inform policy and programs, as well as shape prevention and intervention efforts. One approach to capturing this information is by linking local, state, or federal administrative records, such as those from child welfare, health, social services, education, public safety, and other agencies.

The Child Maltreatment Data Linkages (CMI Data Linkages) project identified five research groups (sites) with experience using linked administrative data to examine child maltreatment incidence and related risk and protective factors. The project supported these sites to enhance their approaches to administrative data linkage through acquiring of new data sources, using new methods, or replicating existing methods. This brief highlights promising practices for sharing and accessing data based on the five sites' experiences. We discuss lessons learned related to four key activities essential to sharing and accessing data: (1) developing agreements for data sharing and use; (2) protecting the data's security, confidentiality, and privacy; (3) securing institutional review board (IRB) and other approvals; and (4) accessing the data. Additional detail can be found in the full report, [Linking Administrative Data to Improve Understanding of Child Maltreatment Incidence and Related Risk and Protective Factors: A Feasibility Study](#).



Promising practices: sharing and accessing data

- Researchers can build trust with data partners by making sure they know the federal, state, and local laws and agency-specific regulations regarding data access and by collaborating with liaisons in public agencies.
- Organizations interested in enhancing data linkages might be able to modify or amend existing data use agreements (DUAs) or research permissions to conduct their work. An IRB that allows organization to use the data for numerous projects and analyses of linked administrative data might also support these types of projects.
- Having a principal investigator (PI) with experience and knowledge of the IRB process may simplify the approval process. Project teams can use this experience to submit thorough materials, resulting in fewer revisions to the submitted IRB package.
- Although DUAs often require stringent data security protocols, research centers working with administrative data might already have such protocols in place. Collaborations with external entities to conduct data linkages are an additional means to ensure the privacy and confidentiality of personally identifiable (PII) information.
- Using publicly available data or data that a variety of users frequently use (where there are established procedures in place to access the data) can simplify data sharing.
- Plans and timelines for projects involving administrative data linkages should build in room for delays, especially related to data acquisition, and identify opportunities to accelerate other activities or use the time to prepare for analyses. ▲

Table 1. CMI Data Linkages Projects

Replicating the Alaska Longitudinal Child Abuse and Neglect Linkage (ALCANLink) methodology Alaska Department of Health and Social Services and Oregon Health Sciences University (ADHSS/OHSU)

The ALCANLink approach used a population-based, mixed-design strategy to integrate two sets of data: (1) those births that were sampled and mothers who subsequently responded to the Pregnancy Risk Assessment Monitoring System survey and (2) child welfare and other administrative data. Alaska partnered with Oregon to replicate this methodology and to estimate and compare the cumulative incidence to first report, screen-in, substantiation, and removals by age 9.

Methods to estimate the community incidence of child maltreatment Children's Data Network and the California Child Welfare Indicators Project (CDN/CCWIP)

This site focused on developing a methodology that used administrative data to estimate the number of children who were victims of abuse or neglect. The site produced upper and lower bounds of estimates that reflected the number of children who the child welfare system identified as victims of abuse or neglect, as well as those who were victims but not identified as such by the system. The site tested the methodology using data from California and explored the potential for using it in other states.

Using hospital data to predict child maltreatment risk Children's Data Network and Rady Children's Hospital-San Diego (CDN/Rady)

This site tested the predictive value of integrating hospital data with vital birth records, statewide child protection records, and vital death records to identify children who might be at an elevated risk of maltreatment. The site focused on validating a statewide predictive risk model by determining the extent to which children identified to be at high risk of maltreatment are also at elevated risk of injury, poor health outcomes, and mortality in childhood. The site used machine-learning methods to train probabilistic algorithms for linking hospital-system data to other administrative data sources. These data linkages aimed to better characterize the demographics and public service trajectories of Rady Children's Hospital patients.

Understanding the effect of the opioid epidemic on child maltreatment Center for Social Sector Analytics and Technology (CSSAT)

This site contributed to the knowledge about the opioid epidemic's potential effects on child maltreatment. Drawing from several data sources across Washington State, this project examined the associations among multiple indicators of child maltreatment, child welfare system involvement, and individual- and community-level risk factors.

Examining child maltreatment reports using linked county-level data University of Alabama School of Social Work (UA-SSW)

This site examined how risk and protective factors relate to child maltreatment reports at the county level across the nation. The site linked county and state data from the National Child Abuse and Neglect Data System to county and state data from the U.S. Census, Bureau of Labor Statistics, Center for Disease Control and Prevention, National Center for Health Statistics, and other sources. The site aimed to explain widely varying state- and county-level maltreatment rates and to develop valid ways to use county-level child maltreatment risk.

Developing agreements for data sharing and use



Sites relied on existing and new agreements with data partners to access data necessary to complete CMI Data

Linkages studies.¹ Sites aiming to access new data (CDN/Rady, ADHSS/OHSU, and CSSAT) needed to identify who had the authority to grant access. These sites also had to determine the appropriate processes for making requests. Understanding the structure of state agencies was important for completing data sharing agreements and identifying whether multiple approvals were required. Some projects (ADHSS/OHSU and CSSAT) worked with more than one state agency and had to identify the approval authority within each. In one site (ADHSS/OHSU), the core team included an advocate in one state agency (Oregon Health Authority). This person was an effective liaison with her own agency and other state agencies that provided data or supported the analysis. In another site (CDN/Rady), the research team had to identify the people with approval authority to share hospital admissions data: a chief administrative officer at the hospital and a transactions officer at the university medical school the hospital is affiliated with.

All sites had DUAs between the principal investigator (PI) or research organization and each separate agency. No sites were required to have multiparty DUAs. Multiparty DUAs can be more cumbersome to establish because they require coordination from multiple agencies.



Provisions in sites' agreements with data partners focused on administrative, technical, and physical safeguards to protect data.

For example, agreements included requirements that data be transmitted and stored securely; that they not be moved, copied, or transmitted without safeguards; that they not be sold; that confidentiality was protected, and no identifying information revealed in any research data sets or publications; that access to data be limited to only those directly involved, and that data breaches be reported as soon as possible. The agreements also specified that facts cited about the

data must be accurate, and that the data could only be used for specified study purposes. Researchers built trust with data partners by demonstrating familiarity with laws and regulations about data access, as well as procedures for protecting data.



Across sites, DUAs contained similar provisions about the use of data, dissemination of results, and procedures for responding to disclosures of information.

The specific details of each DUA varied by agency. For example, several sites' agreements with the state child welfare agency (ADHSS/OHSU, CDN/Rady, and CDN/CCWIP) specified that research must support the missions of public health and child welfare agencies, and that the agency must be consulted about analysis results and dissemination products before any dissemination takes place. The ADHSS/OHSU agreements also specified a "minimum necessary information" policy: researchers must request only the data necessary to answer their research questions. Yet another agreement specified that research staff consult with the data partner (CDN/Rady) about any disclosures that might be required by law, so the data partner could consider how to respond.



The level of flexibility for sites to use acquired data for additional or alternative analysis varied by site, as specified in their DUAs.

For example, CDN/CCWIP, CDN/Rady, and CSSAT have broad DUAs with their child welfare agencies, allowing them to use the data for numerous projects and analyses. This broad license was an asset for their CMI Data Linkages projects because they did not have to reestablish access or permission to use the data through a new DUA. These broad DUAs still have agency review requirements for use of the data even though the sites did not have to reestablish the access or permission. For example, in the CDN/CCWIP and CDN/Rady sites, the DUA between CDN and the California Department of Social Services (CDSS) allows CDSS data to be used for "research purposes specifically ... sanctioned in writing by CDSS." In contrast, the ADHSS/OHSU site had permission to use child welfare data for the CMI Data Linkages project specifically, rather than broad authorization.

Types of provisions commonly included in DUAs

Data be transmitted and stored securely

Safeguards necessary to move, copy, or transmit data

Restrictions on selling data

Confidentiality is protected and no PII revealed in any data sets or publications

Access would be limited to only those directly involved

Data breaches would be reported as soon as possible



Using publicly accessible data simplified and accelerated data acquisition for some sites.

No data-sharing agreement was required for data sets in one site (UA-SSW), but an application process was required to obtain the data from the National Data Archive on Child Abuse and Neglect (NDACAN). Because NDACAN data are de-identified data submitted by states, the data UA-SSW received were already clean. Other sites (ADHSS/OHSU and CSSAT) used some administrative data sources, such as vital records, that a variety of users frequently access. In these sites, states had established procedures for sharing these types of records, which involved a request and standardized transaction rather than a full partnership and approval process. These data sources offer advantages in terms of ease of access, but they also present limitations. Policies on these data sources are subject to change. Though the process is standardized, it can still be hard to access these data sources in most states, especially in a way that allows individual-level linking.

Protecting the security, confidentiality, and privacy of data

To access and use data, sites had to meet security standards established by multiple agencies and institutions, but sites' existing protocols were stringent enough that they did not require adjustments. Sites had standard security protocols in place, which they were able to use to meet CMI Data Linkages project requirements.



Sites' protocols abided by the separation principle— they separated personally identifiable information (PII) from analytic files and used it only to link records.

In two sites (CSSAT and ADHSS/OHSU), research staff did not have any access to PII. An external third party completed the linkages and returned a completed research file via encrypted transfer, with no identifying information. This increased data security because no team members had direct access to the individual-level records. In two other sites (CDN/CCWIP and CDN/Rady), a select group of non-research staff processed PII only on non-networked

computing stations and used it only to link records. These data were not backed up externally, only to specific encrypted devices. After linkages were completed, restricted analytic data sets were stripped of all direct identifiers and created and processed on a secure data server.

In the site using only publicly accessible data from NDACAN (UA-SSW), the research team members still abided by standard security protocols. For example, they used double-password-protected computers in locked offices with encrypted cloud storage. In addition, to prevent the potential for identification of individuals, NDACAN policy does not permit the sharing of county-level data for counties with fewer than 1,000 child maltreatment reports.

Securing IRB and other approvals



Some sites secured rapid IRB approvals or modifications, whereas other sites experienced prolonged delays.

Three sites that aimed to add new data sources to existing data linkages (UA-SSW, CSSAT, and CDN/Rady) submitted IRB amendments or modifications to existing IRB packages. Two sites (UA-SSW and CDN/Rady) completed IRB modifications that were approved relatively quickly. University IRBs approved the modifications. The UA-SSW analysis did not involve individual identifiable data and thus had an easier IRB process.

One site seeking approval of an amendment through a state IRB (CSSAT) faced substantial delays. Because an initial amendment request did not include all the variables the team needed to access, they needed to submit another amendment. The team waited three months for approval of this amendment. An IRB amendment was also required because of a change in personnel for the project—specifically, the person who was linking the records. Finally, additional state and university approvals were required because the home institution of the co-PI changed. In all, IRB processing in this site lasted about six months.

The ADHSS/OHSU site required a new IRB for the project. Team members noted that the process for securing a new IRB approval (from a university) was relatively smooth. For example, the board granted

approval within the expected time frame and did not request substantial new information about the planned approach. Team members attributed the positive experience with this approval process to an advisor's previous experience with the process.

Accessing data

Sites were generally successful in accessing the data needed for their CMI Data Linkages projects. Sites accessed and used 18 of the 20 planned data sources in their analyses.



Sites' experiences accessing data varied by the type of data.

Although some sites had DUAs in place for child welfare data before their projects began, establishing agreements to share and use these data generally required substantial time and effort. Teams also navigated challenging processes to access hospital data (CDN/Rady and CSSAT), which required negotiation with multiple parties or relatively complicated approvals. The CDN/Rady team shared its impression that the process for acquiring hospital data was more cumbersome than for other types of data because private hospital data has not been used for research as much as other types of data, such as vital records. Teams using vital records data generally found these records relatively easy to access because state agencies had established procedures in place to share them.

For the two data sources that sites were not able to access, the types of data differed, and so did the reasons that sites were not able to acquire them. A site seeking statewide data from a prescription management system (CSSAT) was unable to access these data because of problems engaging the data partner, the state's department of health. The site did not have an existing relationship with this data partner and found that communicating with key contacts was difficult to sustain in the context of the COVID-19 pandemic. A site using National Child Abuse and Neglect Data System (NCANDS) data (UA-SSW) was unable to access data for counties with fewer than 1,000 child maltreatment records. The Administration for Children and Families established this threshold to lower the risk that people living in smaller counties could be identified.

Conclusion

The experiences and findings of the CMI Data Linkages sites offer important lessons about the process of using administrative data linkages to study the incidence of child maltreatment and related risk and protective factors. These lessons underscore the potential for these approaches to inform understanding of child maltreatment.

Sites benefited from existing infrastructure and relationships, which took time and effort to establish and maintain. To accomplish their projects, the sites drew on existing relationships with data providers, existing technical expertise, and existing infrastructure. The sites nurtured relationships with data stewards through regular meetings and consideration of the child welfare agency's priorities when conducting research—for example, considering and communicating how the research could help the data providers as well as the site. PIs and co-PIs were seasoned researchers with expertise in administrative data linkage and analysis. Nearly all sites that needed agreements to access child welfare administrative data already had them in place. However, the sites' data linkage projects required substantial effort and resources, particularly if researchers did not have existing infrastructure and experience.

Although sites' existing relationships, expertise, and infrastructure proved helpful in many circumstances, existing relationships with data stewards did not guarantee smooth processes for sharing additional or new data.

The sites needed to adapt to changes in circumstances and address unforeseen challenges that affected their project plans. All sites had to adapt to changes in working conditions, priorities, and partner availability resulting from the COVID-19 pandemic.

Overall, the CMI Data Linkages sites implemented promising practices for sharing and accessing data that enabled them to address high-priority questions about child maltreatment incidence and related risk and protective factors.

Endnotes

¹ For additional information about data use agreements and research approvals by site, see Tables III.1 and III.2 in [the full report](#).

April 2022

OPRE Brief: 2022-107

Project officers: Jenessa Malin and Christine Fortunato

Project Director: Matt Stagner

Disclaimer: The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research, and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services. This report and other reports sponsored by the Office of Planning, Research, and Evaluation are available at www.acf.hhs.gov/opre.

Connect with OPRE

