Introduction

High-quality early care and education (ECE)\(^1\) experiences can improve the health, academic, social, and economic outcomes for young children and their families.\(^2\) These outcomes are more effectively achieved when community partners work together to link data to support high-quality early experiences and services that promote children's optimal development. Having access to current and past data collected on children and families by other local or state agencies, in combination with use of their own program data, can contribute crucial information to support high-quality practices in ECE settings.

In recent years, states have started developing and using early childhood integrated data systems to better support children birth through age five.\(^3\) However, there are fewer efforts supporting local programs to combine their own data with data from other organizations to improve services. Therefore, this brief focuses on understanding the benefits and challenges for local ECE programs in sharing and receiving (i.e., linking) their data with other organizations at the local or state level to improve services for young children and families.

---

\(^1\) In this brief, early care and education programs include child care centers, pre-k programs, Head Start, family child care homes, and other programs that provide care and education to young children, birth to five, prior to school entry.


There are several benefits of using linked data to help strengthen ECE program quality and services, but there are also many challenges that must be overcome throughout the data sharing process. This brief begins by defining the terms *administrative data* and *linked data*. We then describe benefits for ECE programs of using linked data and provide examples of how ECE programs, including Head Start programs, have linked their data to make decisions related to improving program quality and services. We draw upon information gathered from published reports as well as conversations with Head Start grantees, city-level administrators, state-level data systems specialists, state child care administrators, technical assistance providers, and researchers (please refer to the Appendix for more details). Finally, the brief highlights challenges programs encounter when attempting to link data, or to use linked data, and offers strategies to overcome the identified challenges.

**Defining Administrative Data and Linked Data**

In this brief, *administrative data* refers to information programs collect about individual children, families, and staff to deliver program services and meet program, funding, or legal requirements. Generally, programs collect administrative data to determine child/family eligibility for services, monitor staff workload, document services provided, or examine progress children are making. Examples of administrative data may include information on enrollment, staff qualifications, family eligibility, child assessments, child health, mental health and behavior, participation in child welfare, and demographic information. Local ECE programs vary in the types of data they collect and how they use data, often based on different funding requirements (e.g., Head Start, state-funded prekindergarten) and other factors. Programs often keep these data in multiple places or databases.

In this brief, data are considered *linked* when information from two or more separate data systems or databases are shared, connected, combined, or merged. These data systems or databases may be housed in the same program or in two or more separate agencies. Data linking can occur in various ways, ranging from simple to more complex. However, all types of data linking have the potential to provide useful information for ECE program improvement. For example, a relatively simple way to link data is having staff from two different agencies enter data into the same spreadsheet. As another example of linking, multiple ECE programs might enter data into a single, shared database on a website. This may be somewhat complicated to set up but relatively easy for staff to use. Another simple way to link data involves staff at one agency sharing data on families with staff at another agency (e.g., enrollment information from a social services agency; immunization records from a health department; housing status from a community housing provider), so that staff at the second agency can enter this new information into their own database. A third way to link data is to merge two separate datasets into a single data file. This is more complex and requires some way of matching data about an individual in one dataset with data about that individual in another dataset (e.g., matching individual children across files based on name, date of birth, and ethnicity or using a unique identifier if available). Two organizations often engage in this kind of data linking as part of a research study or project. As a final example, several states and communities are developing *integrated data systems* that link data across multiple programs or agencies that serve children and families. This is one of the more complex ways to link data that often requires time and resource investment from multiple agencies. Examples of information that can be linked for ECE program improvement include:

- Child, family, or staff information from different data systems within an ECE program (e.g., child attendance, family characteristics, teacher training, child assessments);

- Other ECE program data (e.g., early intervention, state pre-k, child care);

- Data from health and social service programs (e.g., Temporary Assistance for Needy Families or TANF data from a local social services agency; Medicaid information from the health department; Supplemental Nutrition Assistance Program or SNAP data from a local social services agency);

- Local K-12 education data (e.g., kindergarten entry assessment, attendance, special education placement, 3rd grade reading level); or

- Many types of data in a community or state integrated data system (i.e., immunization, child welfare services, job training for parents, developmental screening information, K-12 education data).
Benefits of Using Linked Data

Programs can improve the services they provide when they are able to link administrative data within their own program as well as share data across agencies. This ability to link data across multiple agencies also supports the development of strong partnerships and early childhood systems at the local or state level. This section outlines important benefits of using linked early childhood administrative data and provides examples of how ECE programs have used linked data to improve their own services and those in their communities.

Inform Staff Support

By linking data from multiple sources, ECE programs can improve the supports they provide for staff and, in turn, better support the needs of young children and families. For example, having ECE data that are linked can inform the kinds of professional development and training needed at the program level. One program that we interviewed linked children's assessment data with their data on staff qualifications and training. By linking these data within their own program, they were able to compare children's performance before and after program staff went through a targeted training on teaching math and science. Linking data can help program directors examine how staff training relates to children's skills so that they can appropriately allocate resources to support continued staff development.

Examine Children's Progress after They Leave the Program

ECE programs are often interested in understanding how children are progressing once they leave the program and move into elementary school. By combining ECE program data with public school data, ECE program leaders can have access to information on how children are performing in later years. One ECE program that we spoke with tracked children's progress through 3rd grade to examine the trajectories of children who attended their program. The ECE program worked with the local school district to gather grade retention and achievement data on their ECE program graduates over a 10-year period. The ECE program leaders were then able to look at trends over time and use the information to support conversations between the ECE program and elementary school, and inform transition supports for children and families. The ECE program and school personnel also used these linked data to help them plan professional development activities for staff. This kind of long-term collaborative data-linking project can allow ECE program leaders to examine trends over time and look at multiple years' of data to identify possible areas for improvement. However, linking and then analyzing administrative data over time is not necessarily the same as conducting a more rigorous program evaluation, and findings from analyses should be interpreted cautiously.

Improve Comprehensive Services for Young Children and Families

ECE staff can have a more comprehensive picture of the families they serve and can better individualize services if they have data on other social service programs in which their children and families are enrolled. Utah, for example, is linking data from multiple programs, including Head Start, health, and social service agencies through a state-level data system. This helps programs verify whether children have received their immunizations and identify if families may need additional support in this area. Additionally, there is interest among ECE program leaders to link their data with data from other social services (e.g., Supplemental Nutrition Assistance Program or SNAP, Temporary Assistance for Needy Families or TANF) to better understand the needs of the families they serve, to inform their community assessments, and to identify additional support services available to families. For instance, a Head Start program we spoke with that linked data with a local social service agency was able to identify if any of their families were eligible for but not receiving TANF. They could

---

then provide TANF enrollment assistance for those families in their program. Through data linkage efforts, ECE programs would have more comprehensive information about the children and families they serve.

Reduce the Burden on Program Staff

Program staff sometimes must enter information into multiple systems or databases, particularly if they receive funding from multiple agencies and need to report data separately for each funding source. For instance, if a Head Start grantee receives funds from the state pre-kindergarten program and the child care subsidy system, program staff may be required to enter child attendance and family income into two different data systems. If program staff linked the data from these multiple systems, then they would only have to enter the information once, reducing their data entry time so they could focus on other program tasks. This would also minimize the chance of making errors when entering the data.

Having electronic access to data about children and families from outside agencies can also reduce the burden on staff. For example, we heard from a Head Start program that their staff has access to child screening and assessment data about some of the children in their program who were previously enrolled in a program that used the same data system. Staff were able to view the previous assessment information in the data system. They used the information to better understand the children's unique needs, which provided more insight as they began to conduct their own assessments. Staff also had less data to gather or enter on these children because it was already captured in the data system.

Reduce the Burden on Families

Families needing assistance in a number of areas (e.g., housing, health care, child care) often have to provide the same information (e.g., children's date of birth, family income, immunization records) to every program from which they receive services. If a Head Start program linked its data with data systems from other community agencies, then staff could know if families applying to their program were eligible for or already enrolled in other programs (e.g., SNAP, TANF). It might also be possible for eligible families to enroll in these other programs without having to go to multiple agencies or provide the same information multiple times. This would not only reduce the burden on families but also make it easier for Head Start staff to provide support services to families.

An ECE supervisor we talked to told us about their state's effort (Colorado) to create a universal intake form used by multiple programs to enroll families. With this universal intake form, a family only has to complete one application to verify their eligibility for various programs and services within the state, including child care. This reduced the burden for families as it helped them quickly identify programs for which they were eligible while also minimizing staff time to verify eligibility. Over time, if children remain enrolled in the program, program staff will likely need to continue to understand the family's circumstances and eligibility, as required by the given program's guidelines.

Inform Strategic Community Planning of Services for Children and Families

Head Start and other ECE programs are a critical part of communities, and many program leaders are actively engaged with state and local partners to strengthen the array of services provided to children and families. Head Start grantees are required to conduct a community needs assessment to guide their services, and it must be reviewed annually and updated if there are significant changes to the information in the assessment. Linking data across multiple programs and services could help Head Start and other leaders better understand community needs and plan services to meet the needs of children and families. Program and community leaders who have a complete picture of which children and families are accessing ECE programs and other services, along with the quality of those programs, can identify service gaps for specific populations (e.g., infants/toddlers, dual language learners, low-income families) and implement funding or policy changes.

5 The universal intake form is available here: [http://www.coloradoofficeofearlychildhood.com](http://www.coloradoofficeofearlychildhood.com). From this website, click on Quality Initiatives and then click on Universal Application.
Challenges and Corresponding Strategies for Linking Data

Linking administrative data, and then using these linked data, can greatly benefit ECE programs by supporting continuous program improvement, but it can be hard to do. This section discusses challenges programs might face when trying to link their own data or link their data with data from other agencies. We also describe strategies ECE programs have used to overcome these challenges.

Capacity to Use Linked Data to Inform Practice

**Challenge:** Program leaders and staff may value using data and want to link data but may not have the capacity within their program (e.g., staff time, technical knowledge, resources) to move forward. Multiple supports are needed to create a strong data culture that would allow staff to link data and use that combined data to guide discussions on program improvement. With a strong data culture, directors and staff use data regularly for planning and decision-making purposes. They also are able to develop the skills to manage, analyze, link and use data. This includes having or hiring staff with the technical skills needed to work with and link to multiple data systems, as well as staff who have experience and training to use linked data to inform practice and program improvement while recognizing its limitations.

**Strategies:**

- **Implement program policies to use data regularly.** For example, one Head start director we spoke with decided to have teachers present their classroom data during routine meetings. This gave staff the opportunity to use data, discuss challenges, brainstorm strategies, and monitor improvements in children’s growth and development.

- **Identify ways to simplify data collection for staff.** For example, staff in one program we talked to were given laptops to make it easier for them to collect and use data regularly.

- **Offer training to help staff appropriately use data.** Programs we spoke with found it helpful to attend trainings or provide mentors to help leaders and staff use linked data. For example, when we spoke with the Head Start National Center on Program Management and Fiscal Operations (PMFO), a Head Start technical assistance provider, they described trainings they offer to Head Start programs to help foster a strong data culture among program leaders and staff. During these trainings, Head Start program management teams work together to understand how to identify, share, and use data to develop continuous improvement plans.

Privacy Issues and Concerns about Sharing Data

**Challenge:** Confidentiality questions can arise when outside agencies request information about children and families from ECE programs. Rules and restrictions about what data ECE programs can legally share are complicated to understand. Program staff may also hesitate to share data if they believe others will use the data to make inappropriate comparisons across children or programs, or inappropriate decisions about individual children or programs. Because of the vulnerability of the families they serve, Head Start grantees may feel a particularly strong responsibility to protect the privacy of family- and child-level data and may have concerns about violating the trust of families who confide in them.

---


7 The resources developed by the PMFO center on Head Start data, including online interactive learning modules, can be found at this link: [http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/operations/data](http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/operations/data).
Additionally, families themselves may have concerns with the ECE program sharing personal information on their family or child with other social service or education agencies. These concerns and questions can make it challenging to link administrative data across programs and services.

**Strategies:**

- **Establish a data governance body.** Some communities have developed data governance bodies to determine rules for using and sharing data. These governance bodies can do several important things to build trust between programs and the families they serve, as well as between multiple partner organizations that are sharing and/or receiving data from each other, to minimize privacy concerns:
  
  ✔ *Include representation from every agency that shares data, as well as family and community stakeholders.* This ensures the views and concerns of every stakeholder are considered when policies are developed or decisions are made. For one Head Start grantee we interviewed, the council included both program and parent representatives who helped think about the data collected and how to appropriately use the data.

  ✔ *Develop data policies and procedures.* These policies and procedures can address privacy concerns related to data collection, access, and use. It is important to include families and staff in the development of data policies. The policies and procedures will provide important guidance to staff and should be reviewed regularly and revised as needed.

  ✔ *Develop parent consent and notification processes.* These processes include communications with families about their rights to share or keep personal data confidential. It is important for program staff to: discuss the data they would like to collect with families; seek feedback and input from families regarding their interests and concerns around data collection, use, and linking; and develop procedures that clearly explain families’ choices and involvement in the collection, use, and linking of data.

  ✔ *Be transparent about how data are used.* For example, a city administrator who oversees early childhood services told us that in their community, researchers and other agencies requesting to link and analyze early childhood data are required to describe to the governance committee exactly how they will use the data and for what purpose before the committee grants permission to use the data.

- **Develop data use agreements.** Programs can create their own data use agreements to specify which data can be used and by whom. These agreements describe the required procedures another program or outside individual must follow when using the program’s data in order to protect the privacy of the children and families. For instance, programs can create data use agreements requiring that anyone who uses their data must store the data in a secure location and cannot publish or release any information that could identify children, families or staff. Data use agreements also often outline policies and procedures for the destruction of data after it is no longer needed. Programs typically develop data use agreements for a specific agency or individual who has requested to use the program data for a particular purpose; alternatively, the requesting agency or individual may already have a data use agreement that the program can modify for their purposes.

### Managing Multiple Data Collection Systems

**Challenge:** A particular challenge in using linked data is determining how to connect data from multiple databases and across multiple programs. Many ECE programs receive funds from multiple agencies (e.g., Head Start, state pre-k, child care subsidies) that may have different reporting requirements. As a result, much of the data about children, families and staff participating in ECE programs are stored in different data systems. The data may be collected differently in each data

---

* In many cases, it is not necessary for a program to share personally identifiable information (PII) in order to link data with another organization. A program should only share PII on children and families if absolutely necessary. As an alternative to sharing PII, the program can share de-identified data using unique identification numbers that are assigned to each child or family. See the discussion of unique identification numbers below for more information.

* For more information on communicating with families about data, see another resource in this series, *Data Direction #2: Discussing Data with Families.*
system (e.g., monthly family income reported in one system and annual family income reported in another), as well as defined differently across the data systems (e.g., infants may be defined as birth to 11 months in one dataset and birth to 12 months in another).

**Strategies:**

- **Assign unique identification numbers.** One strategy to link data is to have a unique identifier (UID). A UID is a single, unduplicated number that is assigned to and remains with a child, family member or staff member throughout participation in early childhood, K-12, and other public programs. Having UIDs for the children, families, or staff in datasets allows information to be linked between two or more datasets because there would be a way to match individuals across various datasets. Some programs we spoke with used UIDs generated from state or local school district data systems. Programs also matched records based on unique information (e.g., name, date of birth, or race/ethnicity) instead of using UIDs, but this process can take a long time and may not yield a complete match of individuals across datasets. A benefit to using UIDs instead of matching records on demographic information is that it reduces the need to share additional personally identifiable information on children and families because UIDs are not connected to any identifiable data on the child or family. For example, instead of sharing data such as name, address, or date of birth, programs can use UIDs to match children and families. This helps protect the privacy of children and families in their programs. However, even when using UIDs, it is important to ensure that data linking procedures and protocols are in place and followed carefully so that the confidentiality of sensitive information is protected.

- **Coordinate data collection systems and practices.** A larger Head State grantee we spoke with invested in developing technology to merge the various data systems used by their programs. Leaders also worked to align data definitions to make it easier to link data across programs (e.g., developed a data codebook so that an infant was always defined as birth through age 11 months) and generally encouraged the development of an integrated statewide early childhood system that could reduce the duplication of data across multiple funding streams. Efforts at the federal level are also helping support better coordination of data collection standards and practices. For example, the U.S. Department of Education's *Common Education Data Standards* (CEDS) initiative is developing voluntary, common data standards for programs serving children from preschool through high school and beyond. The Administration for Children and Families in the U.S. Department of Health and Human Services also developed the complementary INQUIRE Data Toolkit to specifically support effective data collection and data standards for ECE programs.

---

10 To learn more about the Common Education Data Standards, please visit: [https://ceds.ed.gov/](https://ceds.ed.gov/).

Capacity of Existing Data Collection Systems

**Challenge:** Many programs use existing software from vendors to collect, store and report their data to meet basic reporting requirements. These software programs may have limited flexibility in generating other reports or customizing the databases to meet individual program needs to link or report data. For example, a Head Start Director we spoke with wanted to know which children in her program missed more than ten days but was only able to determine this by examining the raw data and counting absences for individual children (rather than creating a report). Another program did not have an existing data field to track parent consent for data collection. These limitations can make it hard for a program to use or share data.

**Strategies:**

- *Work with vendors to customize or upgrade databases.* Some ECE programs have worked with vendors to make changes that facilitate linking data. One administrator told us about working with multiple vendors to modify their data systems so that they could combine child assessment data with other program-level data. As another example, a state-level data manager we spoke with negotiated with a vendor on behalf of Head Start grantees to identify which data fields could be used for a parent opt-out indicator for parents that choose not to share their child’s data with the state’s data system. It may be easier and more cost effective to negotiate with vendors when programs work together or with a state agency to collectively request changes.

- *Analyze data using other software.* We heard from a Head Start program director that she regularly uses Microsoft Excel to help her analyze data because of limitations of the existing data system. She uses the software to analyze, merge, and create charts to display and examine the data. While this process can be time consuming and requires expertise about data analysis and other data software, it allows programs to analyze their data in ways that might not be possible within the primary data system.

**Moving Forward**

Understanding the power of linked data to inform program improvement is an important first step in using linked data. Once program leaders are interested, though, they often must overcome several challenges.

The information in this brief highlights the particular issues that local ECE programs face when linking data and offers suggestions for overcoming challenges. As data use and linking efforts grow, especially at the state and federal levels, it is important to maintain a focus on local programs’ use of data. How can local ECE programs collect data well so that it can be useful for them and included, as appropriate, in local and state data linking efforts? How can local programs best use their data and link with other data to improve their program and support children and families?

It is equally important to understand the limitations of administrative data and the cautions when using administrative data to inform program decisions. For example, addressing particular questions such as program effectiveness or the impact of a program on children’s later outcomes would require a more sophisticated evaluation, likely including the collection of additional data.

Local program directors, technical assistance providers, community and state leaders, and families should work together to discuss their shared vision for strong early childhood services and the role of linked administrative data in helping programs and communities realize their vision. We hope these stakeholders will work collaboratively to overcome the challenges identified in this brief. Though this brief offers some strategies to overcome particular challenges, future efforts may provide resolutions that are more permanent. Software packages and vendors, for example, could offer more
flexibility to address local program needs and interests. State early childhood integrated data efforts could support a more coordinated set of data definitions that make it easier for local programs to report information consistently.

As part of the Building Capacity to Use Linked Data project, we have developed a series of resources to support ECE programs in linking and using linked data, both within their own program and with other local and state agencies. They include a set of case studies highlighting examples of programs that have successfully shared and linked their data; a set of tips and guidance documents for addressing common issues that arise when linking data; a data planning guide for program administrators to examine their own progress in linking data; and a compendium of existing resources on these topics.
Appendix: Methodology

This brief was developed through a review of existing literature, discussions with expert consultants for the Building Capacity to Use Linked Data (BCULD) project, and conversations with early care and education (ECE) stakeholders at the local, state, and federal levels. We first conducted a search of the literature using various databases to identify reports, fact sheets, webinars, case studies, and guides/tools published within the last 10 years that discuss opportunities, challenges, strategies, or recommendations for using and linking data in Head Start and other ECE programs. We used a broad list of search terms, including linking early childhood data, Head Start and linking data, data governance, data quality, data use, data security, and data management. From this initial list, we then reviewed 25 resources that focused on benefits, challenges, or recommendations about linking and using linked data. We also conducted a targeted search of the national Head Start technical assistance centers to find resources specific to Head Start data system linkages.

Next, the BCULD team solicited feedback from expert panel members about their experiences and ideas on the benefits and challenges of using linked data. The BCULD team also asked the experts to identify other individuals with similar experience. Each expert recommended up to three stakeholders that could provide a local-, state-, or federal-level perspective about the use of linked data in Head Start and other ECE programs. The BCULD team selected nine stakeholders to participate in informal phone conversations to share their thoughts and experiences on these issues. This brief summarizes the findings from relevant existing resources, conversations with the expert panel members, and selected stakeholders on benefits and challenges of linking and using linked administrative data.

---

1 Information sources included: Google Scholar, Elton B. Stephens Company Information Services (EBSCO), Research Connections, Quality Initiatives Research and Evaluation Consortium (INQUIRE), BUILD Initiative, Early Childhood Data Collaborative (ECDC), Data Quality Campaign (DQC), The National Workforce Registry Alliance (The Alliance), Common Education Data Standards (CEDS), Quality Rating and Improvement System (QRIS) Learning Network, National Center on Child Care Data and Technology (NCDT), Privacy Technical Assistance Center (PTAC), State Longitudinal Data Systems (SLDS), and The Center for IDEA Early Childhood Data Systems (DaSy).
Acknowledgments

Child Trends is grateful for the financial support of the Building Capacity to Use Linked Early Childhood Administrative Data project, provided by the Office of the Assistant Secretary for Planning and Evaluation, in partnership with the Office of Planning, Research and Evaluation, in the U.S. Department of Health and Human Services. The goal of the project is to help early care and education programs build their capacity to link and use linked administrative data for program improvement and research purposes.

We would like to thank the federal staff who provided importance guidance throughout this project: Lindsey Hutchison and Kimberly Burgess, Office of the Assistant Secretary for Planning and Evaluation; Christine Fortunato, Office of Planning, Research and Evaluation; and Theresa Rowley, Office of Head Start.

We also appreciate the contributions of the project’s expert panel, whose members dedicated time and provided advice to strengthen the resources developed as part of this project. Expert panel members include: Leigh Bolick, South Carolina Department of Social Services; Craig Detweiller, Georgia Department of Early Care and Learning; Emmalie Dropkin, National Head Start Association; John Fantuzzo, Penn Child Research Center, University of Pennsylvania; Vanessa Rich, Chicago Department of Family and Support Services; and Amanda Schwartz, National Center on Health, Amanda Schwartz Consulting.

We extend a special thank you to the early care and education program staff who provided examples of using linked data, explained the opportunities and challenges in linking data, described how they linked data and how it helped inform continuous quality improvement, and talked with us about their programs’ data practices.


This report is in the public domain. Permission to reproduce this document is not necessary.

This and other reports sponsored by the Office of the Assistant Secretary for Planning and Evaluation are available at https://aspe.hhs.gov/reports and other products from the Building Capacity to Use Linked Data project are available at https://aspe.hhs.gov/building-capacity-use-linked-data-program-improvement-and-research-initiatives.

© Copyright 2016 by Child Trends, Inc.