

Administrative Data for Action in the Social Sector

Christina Bath Collosi
July 30, 2018

CONTENTS

- ADMINISTRATIVE DATA BACKGROUND** 03
 - Integrated Data Systems 04
 - IDS: Real-Time and Near Real-Time Planning and Intervention 04

- EARLY CHILDHOOD DEVELOPMENT** 04
 - The Science of Early Brain Development 04
 - School Readiness: Role of Public Benefits and Supports 05

- UNITED STATES: ADMINISTRATIVE DATA SHARING LANDSCAPE** 06
 - The Role of Governors 06
 - Federal Support for Administrative Data 06

- UNITED STATES: EXAMPLES OF INTEGRATED DATA SYSTEMS** 07
 - Allegheny County Data Warehouse 07
 - Philadelphia Quality Preschool Expansion 08
 - New York City: One-City Service 08
 - Florida Longitudinal Data System 9

- UNITED STATES: EXAMPLES OF FAMILY-CENTERED SYSTEMS DESIGN** 11
 - Pennsylvania Integrated Eligibility System for Public Benefits 11
 - Michigan Bridges 11
 - California’s Enrollment Landscape 12

- UNITED STATES: BIRTH MATCH REAL-TIME USE OF INTEGRATED ADMINISTRATIVE DATA** 12
 - New York City 12
 - Maryland 13
 - Michigan 13
 - California 14

- GLOBAL EXAMPLES: ADMINISTRATIVE DATA AND INTEGRATED DATA SYSTEMS** 14
 - Nordic Countries 14
 - Australia 14
 - Canada 15

- CONCLUSION** 15

- REFERENCES** 16

ADMINISTRATIVE DATA BACKGROUND

Over the past 20 years, the requirements for government transparency and accountability have steadily increased. To meet such requirements, both federal and state agencies have started adopting new technological and methodological tools, resulting in large-scale administrative data sets. Administrative data is defined as information about individual children, families, and/or providers of early care and education and other family benefits that are collected and maintained as part of the operation of government programs (*Lin, Maxwell, & Forry, 2017*). Unlike survey data, which is mainly a collection of information from population samples and may prove difficult to re-collect, administrative data represents official information generated from routine government tasks and services. Examples of administrative data sets include birth records, educational records, and participation in public service programs. This data possesses the potential to unlock answers to pressing social questions.

Administrative data can be an incredibly powerful tool for program administrators, policymakers, and advocates in their field of work and for improving the lives of vulnerable children and families. From documenting program outcomes and identifying target population characteristics to assessing policy effects, using data is crucial in the efforts to help children and their families thrive. Data can be highlighted to signify the importance of both the problem and the solution. It provides the primary evidence to make the need case for a program, service, or policy by communicating the characteristics of a population or the effectiveness of existing supports.

Furthermore, while administrative data sets have long been studied by research institutions to inform policy and future practice, combining data sets from multiple sources into an integrated data system (IDS) has made it possible to link and use this data in near real time for better provision of services. An IDS can make it possible to use predictive analytics to target service needs for actual individuals or to develop a profile by common characteristics at or near the time of need.

Receiving the best supports possible at the time of need is especially critical in early childhood. Opportunity equity, and lack thereof, among children and families in California and the United States is well known. There are many efforts underway to highlight and narrow the opportunity gap of children who start kindergarten behind their peers and then stay behind for their entire educational experience. Some progress has been made toward closing this “school readiness gap.” Data shows the gap has narrowed modestly from 1998 to 2010, particularly between high- and low-income students and between White and Hispanic students (*Reardon, 2016*). The same is not true when comparing White and Black students. As recently as April of 2018, U.S. Department of Education’s Office of Civil Rights released data that finds that Black students in the United States continue to be restrained, suspended, and arrested at higher rates than all other students. Research overwhelmingly shows that the opportunity gap presents itself far before children reach the K-12 education system. Children born prematurely, cared for in the foster system, residing in distressed neighborhoods, and speaking languages other than English are more likely to enter Kindergarten less ready than their peers—and most never catch up.

How can the use of integrated administrative data be part of a solution to create opportunity equity among all children? How can integrated data inform real-time or near real-time supports for children and families with the goal of overall health and well-being of those individuals? Why is early childhood a critical time period to prioritize in the promotion of IDS and related program planning, service improvements, and advocacy? What are the obstacles and privacy concerns in linking and utilizing administrative data? What cities, states, and countries are leading the way and what are they learning? In this report, we will address these questions with the ultimate goal of promoting family-centered systems design that leverages the use of administrative data and results in improved child outcomes.

Due to both legal and perceived (but not actual) protections for privacy in the United States, this paper explores the use of IDS in the US as well as in other countries with far more robust IDS.

Additionally, the case studies highlight those places that have prioritized IDS for the purpose of improving education outcomes directly and indirectly, for example, to improve root causes of opportunity gaps such as unstable home environments, lack of access to quality early learning opportunities, and homelessness. Lastly, whenever possible, this paper focuses on examples of administrative data use during the earliest years of life and that links to strong school readiness predictive factors, due to the critical role that early brain development plays on lifelong outcomes.

Integrated Data Systems

Most integrated data systems link a discrete administrative data set for a specific purpose and specified period of time. They link administrative records from multiple agencies to give a broader view of social problems and policy solutions and can provide the information that enables a continuous knowledge-to-practice development cycle. Administrative data captures information about people across their life course while protecting confidentiality. Data collection starts with birth and is gathered through early development, schooling, socialization, transition to adulthood, and adulthood. For some populations of adults, additional information is gathered depending on their characteristics and social supports and systems they may have entered. This can include people who experience homelessness; participate in public benefit programs such as Women, Infant, and Children (WIC); people with certain disabilities; and those who have been incarcerated. As people age, additional information is recorded; the final piece of additional information is the death certificate.

Integrated Data Systems store these records, link them together, and identify them for research purposes. The use of linked administrative data allows policy analysts, program evaluators, and social innovators to test new social program ideas at a lower cost and higher speed than traditional evaluation approaches. The Minnesota Early Childhood Longitudinal Data System (ECLDS) is one example of a robust integrated data system with data collected by the Department of Education, the Department of Human Services and the Department of Health. This integrated data can then be searched online at the population level; any information that could be used to identify individuals has been removed.

IDS: Real-Time and Near Real-Time Planning and Intervention

While most IDS focus on informing future programs, services, and policy at a population characteristic level, some are used to determine eligibility for public services and inform real-time or near real-time service. They can improve the ability of service providers to be efficient, effective, and/or to ensure the safety of individuals. These IDS retain specific personally identifiable information (PII). In most cases, this requires active consent of the person being identified, including the purpose of the consent. In other cases, the PII may not be used, but the characteristics of the population are so targeted that near real-time interventions can be focused on people with those same characteristics. In rare cases, administrative data is used in real time due to a “greater public benefit” or safety of individuals that outweighs individual privacy protections.

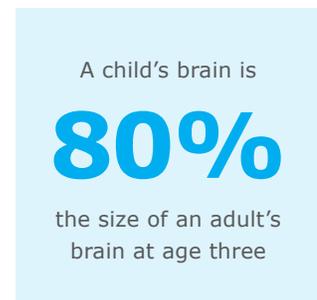
EARLY CHILDHOOD DEVELOPMENT

A child develops in a highly integrated fashion where his or her genetic makeup combined with health, home, caregiving, and environment all contribute to the trajectory of that child’s outcomes. In other words, in the nature versus nurture debate, they both win.

The Science of Early Brain Development

The brain is composed of billions of highly integrated sets of neural circuits (i.e. brain connections) that are wired under the influences of genetics, environment, and experience (*Center on the Developing Child at Harvard University, 2007*). A child’s genes determine when neural circuits are formed, but environment and experiences shape how that formation unfolds. For example, a healthy and loving attachment to a primary caregiver contributes to healthy brain development. In fact, babies appear to be so strongly motivated and prepared to develop attachments to one or more caregivers that, given the opportunity to interact regularly with a responsive caregiver, he or she will develop an emotional tie to that person (*National Research Council, Institute of Medicine, 2000*).

The raw number of actual brain cells (neurons) peaks before birth, but physical brain size increases rapidly in the first five years of life. A child's brain is 80% the size of an adult's brain at age three and 90% at age five. This growth is due mainly to the structure of neurons, which physically resemble miniature trees with a trunk (axon) and hundreds of growing branches (dendrites). Early experiences determine whether a child's developing brain architecture provides a strong or weak foundation for all future learning, behavior, and health (*Center on the Developing Child at Harvard University, 2007*). Striking disparities in what children know and can do are established by kindergarten. These differences are strongly associated with social and economic circumstances and are predictive of subsequent academic performance (*National Research Council, Institute of Medicine, 2000*).



As a person ages, the brain becomes more specialized to assume increasingly complex functions. This leads it to become less capable of reorganizing and adapting to new or unexpected challenges. The basic principles of neuroscience illustrate that providing favorable conditions for healthy development in early childhood is likely to be more effective than treating problems at a later age (*Center on the Developing Child at Harvard University, 2007*). In the past forty years, a compelling body of research has demonstrated that programs aimed at improving young children's health and development have resulted in positive long-term childhood outcomes.

School Readiness: Role of Public Benefits and Supports

It is important to understand the role that public programs play in the early brain development and school readiness of young children—and therefore in children's long-term outcomes.

In 1997, the National Education Goals Panel identified three high-priority objectives that reflect important early supports for school readiness. These have continued as a platform for school readiness. As stated in the Panel's Special Early Childhood Report:

1. Children should receive the nutrition, physical activity, and health care they need to arrive at school with healthy minds and bodies and to maintain mental alertness. To this end, the number of low-birthweight babies should be significantly reduced through enhanced prenatal care.
2. All children should have access to high quality and developmentally appropriate preschool programs that help prepare them for school.
3. Every parent in the United States should be their child's first teacher and devote time each day to helping his or her preschool child learn. To accomplish this, parents should have access to the training and support they need.

For some children, especially those who are born to college-educated and financially secure parents, these three high priority objectives are supported through resources from the family. For many others, families rely on public safety net programs to support their children in these aspects of school readiness. In promoting equitable outcomes for children, it is critical that those who rely on public benefits are able to access and navigate quality services in ways that meet their needs, so that these services can promote optimal school readiness.

How is early childhood a unique window for the use of integrated administrative data? By using research-based predictive data points, it is possible to link administrative data and provide more targeted and effective support while children are still at an early stage of development. Not only does this have the possibility of being more effective due to the malleability of the brain at early ages, but it is also more cost effective to provide early support as opposed to later interventions.

UNITED STATES: ADMINISTRATIVE DATA SHARING LANDSCAPE

The Role of Governors

Governors are in a unique position to advance or restrict the ability of states to coordinate state data across departments and counties. This is done through administrative and legislative action as well as in the communicated priorities of the administration. Ultimately, governors are responsible for being good stewards of taxpayer money and promoting the well-being of their residents. In September 2016, the National Governors Association Center for Best Practices convened a group of state and local leaders and discussed the ways in which shared administrative data can enhance the effectiveness and efficiency of public programs (*S. McGroder, 2016*).

At the convening, several lessons were shared among leaders of states who were coordinating administrative data across agencies to improve services for residents. Those lessons include:

1. Start small and focused and build work from a track record of successes.
2. Focus on a specific issue, for example, Indiana's high infant mortality rate.
3. Alternatively, focus on a low-risk project and then expand once there is success and infrastructure in place.
4. Decide what type of data sharing is required, such as whether data can be de-identified to achieve the desired outcome, and then establish legal data sharing permissions, such as MOUs, that outline privacy and use protocols. For example, MOUs may specify that the data will be used to improve educational outcomes.
5. Establish who owns the data and how it will be governed. For example, is there an intermediary organization outside government that holds the shared administrative data, or is there an office of the government responsible for the privacy and use of administrative data.

Since the September 2016 convening, there has been considerable advancement of integrated data systems and the use of administrative data within public systems. In February 2018, the Pew Charitable Trusts released a report titled "How States Use Data to Inform Decisions: A national review of the use of administrative data to improve state decision-making." The report outlines four ways that states are using administrative data, including: to craft policy responses to complex problems, to improve service delivery, to manage existing resources, and to examine policy and program effectiveness. The report also cites effective leadership as the greatest supporting factor for administrative data use and staffing as the greatest barrier.

Federal Support for Administrative Data

Congress and President Obama enacted the Evidence-Based Policymaking Commission Act of 2016, creating the Commission on Evidence-Based Policymaking, to understand the barriers and make recommendations that would improve the ability to utilize existing administrative data for evidence-based policy making.

The Commission concluded its eighteen-month effort with a September 2017 report, "The Promise of Evidence-Based Policymaking." It discusses the role of evidence building, or using data to better understand the characteristics, behavior, or needs of groups or individuals. The Commission's concluding report contains findings and recommendations in four main areas:

1. Secure and restricted access to confidential data.
2. Enhancing privacy protections for evidence building.
3. Modernizing America's data infrastructure.
4. Strengthening the evidence-building capacity of the federal government.

One of the Commissioners that contributed to this work is local Bay Area resident and UC Berkeley Professor Hilary Hoynes, who teaches Economics and Public Policy and holds the Haas Distinguished Chair in Economic Disparities at the University of California Berkeley. Professor Hoynes has a special focus on poverty and inequality; children, youth and families; and government.

UNITED STATES: EXAMPLES OF INTEGRATED DATA SYSTEMS

Allegheny County Data Warehouse

Allegheny County Department of Human Services (DHS) and its public school partners in Pennsylvania are considered pioneers in the United States for utilizing IDS with the goal of improving educational outcomes for currently enrolled students. After an eighteen-month development process, Pittsburgh Public Schools (PPS) signed a memorandum of understanding with DHS for an “action research” project which provides for DHS and PPS to identify indicators of academic and behavioral success and deficits, analyze the data, and implement strategies and interventions for the purpose of delivering services to students in need (*Fraser, 2013*).

The precursors of this work included the consolidation of DHS from four separate departments into one. Responding to this, a group of local foundation leaders were convened by the Richard King Mellon Foundation. From this convening, a funders group called the Human Services Integration Fund (HSIF) was created under the direction of the Pittsburgh Foundation. HSIF fosters administrative integration and supports innovations in technology, research, and programming that would be difficult or impossible to accomplish with public sector dollars. From the original group of eight foundations, HSIF has more than doubled in size; its member foundations have contributed more than \$12 million to advance the social sector (*Human Services Integration Fund, 2013*).

Central to the successful partnership between DHS and PPS was the creation of a Data Warehouse, a central repository of social services data that now includes information from 19 internal and external data sources representing 29 different human service program areas. This Data Warehouse was initially developed by Deloitte Consulting for \$2.8 million. It transitioned the previous DHS data infrastructure (four separate departments, each with their own database) into one central repository with unique client records. PPS then transfers school records, with personally identifiable information, into this repository. Case workers can view educational data, including things such as grades, attendance, and disciplinary concerns, in real time. Pennsylvania now requires child welfare case workers to complete an education screen for each individual child in their care. In Allegheny County, caseworkers have this information readily available online and integrated with their case files (*Fraser, 2013*).

The Data Warehouse is so robust and useful in real time that it has attracted significant additional funding to support its operations and further investigation and just-in-time intervention. Examples of this include:

1. It was a recipient of a U.S. Department of Health and Human Services Children’s Bureau grant to develop ways to improve educational stability and permanency of children in the child welfare system;
2. The Annie E. Casey Foundation supported additional education stability work; and
3. The John D. and Catherine T. MacArthur Foundation provided funding to enhance the ability of the Data Warehouse to examine the relationship between public housing, education, and child welfare.

The focus on housing has also created opportunities to better understand and find common definitions for terms such as “homelessness.” Prior to the Data Warehouse, the public education system had a broader definition of homelessness than the public housing system, resulting in significant gaps in knowledge, services, and the elevation of priorities when triaging families.

One way that the information in the Data Warehouse is used in near real time occurs through a portal called Client View. This is an individual-level “view” of a client that takes all of the integrated administrative data and connects it together for case managers. Client View provides information about DHS clients and their service involvement, including involvement across multiple DHS program offices. Demographic and education information (as appropriate) is also available, as well as service plans and assessments. Case notes, medications, and addiction treatment information are not available in Client View. Case managers and contracted suppliers utilize Client View to encourage benefit coordination and keep up with benefit progression. District case managers can rapidly audit every one of the public services a family is receiving before making a home visit (*The Annie E. Casey Foundation, 2017*).

Philadelphia Quality Preschool Expansion

In Philadelphia, Pennsylvania, the sixth most populous city in the United States, 37% of children under the age of eighteen live below the federal poverty line. Of these children, half live in deep poverty, which is defined as 50% below the federal poverty line. Leveraging early childhood research, a citywide “soda tax,” which adds a 1.5 cent tax per ounce of a sweetened beverage purchased, took effect on January 1, 2017. The primary purpose of the tax is to fund universal preschool. Philadelphia mayor Jim Kenney proposed the tax, which was approved by the City Council in June of 2016. After nearly two years in litigation regarding the legality of the tax, and \$56 million reserved during this litigation, the Pennsylvania Supreme Court upheld the controversial tax on July 18, 2018.

37%

of children under the age of eighteen live below the federal poverty line

To guide the incremental roll-out of quality preschool expansion, the city used an action-research approach to inform the expansion in real time. This was done through partnership with Penn Child Research Center and the city’s administrative data under the guidance of a newly established Universal Pre-K Commission. De-identified existing administrative data were analyzed for instances of early childhood risks, using research-based indicators, and then the city was geocoded by children with multiple risks. Examples of risks included homelessness, maltreatment, and low birthweight. Geocoding was also done, and updated regularly, to assess capacity of high quality preschool. Using this approach, the city was able to strategically plan and execute where to create or expand capacity based on where there was the most need and where the children had experienced two or more risks. Using the geocoded maps of risk analysis and quality preschool capacity, 23 neighborhoods were identified as being high risk and low or no quality preschool capacity. These neighborhoods were then targeted for the first phase of quality preschool expansion. (*LeBoeuf, 2017*).

Philadelphia is a network member of Actionable Intelligence for Social Policy (AISP), which is an initiative funded by the John D. and Catherine T. MacArthur Foundation. AISP aims to improve the quality of education, health and human service agencies’ policies and practices through the use of integrated data systems. It is led by University of Pennsylvania Professors Dennis Culhane (School of Social Policy and Practice) and John Fantuzzo (Graduate School of Education). The initiative launched in 2008 and has since supported several municipalities to improve policies and practice through the use of effective integrated data systems. AISP has documented several case studies, some of which are highlighted in this paper, and is now formalizing and supporting a network of developing IDS sites.

New York City: One-City Services

In 2002, a few New York City officials set out to coordinate interagency services to more efficiently meet residents’ needs, especially those that rely on services provided by multiple agencies. This “One-City” venture included the Department of Homeless Services (DHS), Human Resources Administration (HRA), Administration for Children’s Services (ACS), Department of Probation (DOP), Department of Youth and Community Development (DYCD), Department of Correction (DOC), NYC Housing Authority (NYCHA), the Department for the Aging (DFTA), and Housing Preservation and Development (HPD) (*Kitzmilller, 2014*).

There were three main outcomes of this administrative data sharing and cross-agency coordination.

1. The benefit dispersal system became more effective.
2. Families had case reviews to coordinate the benefits and programs of multiple city departments, so that they received a coordinated network of support.
3. The agency staffs involved with this coordination reported that sharing administrative data and doing case coordination meant that they were better able to understand each family's needs and barriers.

In 2008, the work was furthered through an executive order from the Mayor that the city's ability to make data-driven decisions needed to be effectively supported to implement data-driven policies and the delivery of efficient services and programs to those most in need. New York City's Center for Innovation through Data Intelligence (CIDI) was established in 2011 as a research and policy center. It is located in the Office of the Mayor of the City of New York and reports to the Deputy Mayor for Health and Human Services.

CIDI has not only analyzed administrative data to make policy recommendations, but it has also layered the data with predictive modeling to make recommendations for families. For example, CIDI partnered with the Department of Homelessness to predict whether a family needing housing would likely need a short or long stay, which helps them place the family in the best housing solution.

Florida Longitudinal Data System

Due to the decentralized nature of education in the United States, the prospect of national educational data records is nonexistent, at least in the foreseeable future. However, several states and school districts have established systems in which longitudinal studies can be used to follow students and match them to schools and, in some cases, their teachers. The three states boasting the longest-standing state-level student longitudinal data systems widely available to researchers are Texas, North Carolina, and Florida. Florida, within its Education Data Warehouse (EDW), has maintained statewide records of student test scores that are longitudinally comparable (initially in grades 4, 5, 8, and 10, and soon thereafter expanded to grades 3 through 10) since the 1997-98 school year. The EDW also includes information on high school graduation, student attendance, grade retention, school assignments, disability, disciplinary infractions, home language, school transcripts, immigrant status, race/ethnicity, gender, and a measure of student poverty, along with other measures. Florida achieved this feat by assigning a unique identification number that remains with the student no matter which school district he or she attends within the state (*Figlio et al., 2015*).

Since the assessment year 2001-02, Florida has linked teachers with their students in each class, allowing the matching of staff databases that include both the teacher and student credentials. This student-level longitudinal data has been used to study a wide range of research questions. For example, it has helped identify the effects of changing school-level accountability on student test performance in public schools, as well as the degree to which such accountability policies affect student sorting across schools (*Figlio et al., 2015*).

The Florida Department of Education, in collaboration with other Florida state agencies, has successfully linked K-12 data with a variety of post-secondary data sources, including employment and earnings records, military service and criminal justice records, and information on post-secondary education and training. These matched data sets are known as the Florida Education and Training Placement Information Program (FETPIP). The FETPIP has followed students in Florida, starting with the 1996-97 high school graduation cohort (*Figlio et al., 2015*).

This has made earnings data available for individuals as late as their early mid-career. FETPIP data has been linked to K-12 education records several times to arrive at empirical research in education. It has been used to compare labor market outcomes of students who barely graduated to those who only just failed these examinations in order to identify the degree to which a high school diploma signals value in the labor market. In a particularly inventive use of this data, the students' test scores are related to their teachers, creating a Value-Added Model that measures how individual

teachers “add value” to their students’ education. This has been used in part to observe whether teachers with higher value-added scores commanded higher labor market earnings when they left public school teaching (*Figlio et al., 2015*).

In recent years, the data has been used to track Florida students not just forward into the labor market, but also backward to early childhood program participation and birth records:

- The Florida Departments of Education and Health have matched the birth records of all children born in Florida between 1992 and 2002 to the school records observed by the Department of Education. To date, over 1.6 million children’s administrative records have been matched between agencies for these research purposes.
- These matched data sets have created a new avenue to study questions ranging from the consequences of poor neonatal health and early interventions for autism spectrum disorders on children’s educational outcomes, to the spillover effect that having disabled siblings could have on children’s cognitive development.
- The earliest of these birth cohorts are, at the time of writing, old enough to have graduated from high school. So, following children born in Florida through their K-12 schooling career shall soon become a reality thanks to the FETPIP data. They can be tracked into post-secondary education, the military, the labor market, and the criminal justice system (*Figlio et al., 2015*).

With the emergence of population-level data, many new research programs can be launched in Florida at a fraction of the cost of purpose-built survey-based databases. Research programs that previously seemed impossible or unfeasible to undertake are now attainable with the generation of high-quality administrative data. The linking of school records to birth and early childhood records and labor market data is allowing researchers to further study certain questions, specifically applicable to the United States, that are otherwise difficult to research (*Figlio et al., 2015*).

For example, Florida’s linked birth and school records allowed the study of the effects of birthweight on cognitive development from kindergarten through a child’s academic career. Using twin studies and single birth comparisons, researchers found that poor neonatal health and low birthweight has significant negative effects that are demonstrated at kindergarten and are persistent throughout schooling. Although early interventions can remediate some of the gap, the effect is largely consistent from kindergarten all the way through a child’s educational experience (*Figlio et al., 2014*). Although there has been significant research showing the effects of low birthweight on cognitive results in schooling, there had never been a study of the variants, such as quality of schooling. The Figlio 2014 study researched 1.3 million single births and 15,000 twin births and demonstrated that, although improved school quality has been associated with improved outcomes for all children, “they do not reduce the gaps generated by poor neonatal health.”

It is important to note, however, that the linked data for children in Florida is only available for relatively recently-born children and there is no information about parents, except what is on a child’s birth records. In addition to this, the United States is an open economy with a high degree of mobility across states and, therefore, it is natural that there would be considerable attrition from even the most complete set of records. For instance, the birth and school record-linked data include only 81% of the children born in Florida between 1992 and 2002. The remaining 19% children either left the state of Florida or were enrolled in private schools and thus were never included in the public school records. While careful scrutiny of other data sources has made it evident that virtually all of the children’s records were matched between these data sets, it still highlights the shortcomings of matched birth and school records in the United States. Nonetheless, the data from Florida highlights some of the limitless possible population-level research programs that matched administrative data can facilitate in the United States and around the world in years to come. (*Figlio et al., 2015*)

In Florida, research using administrative data has also led to substantial policy changes. For example, the school voucher policy in Florida (and in several other states in the United States) has been influenced by research that documents the effects of school vouchers on selection, participation, and competition in various inter-school events. Florida’s administrative data has also significantly

impacted state compensation policies regarding the accumulation of teacher credentials, as researchers were able to identify the degree to which the attainment of credentials influenced teacher performance in the classroom.

In another example, Washington State changed its teacher retention policies after research utilizing twenty years of administrative data from personnel records showed the relatively limited relationship between teacher experience and value-added performance for students in the classroom. The focal question for this work was “What variables consistently explain teachers’ retention and mobility patterns in Washington state?” The two main populations investigated include all teachers statewide (across all years of experience), and beginning teachers (those with less than one year of experience). Research using administrative data in education has not only led to scholarly advancements, but has also initiated positive results in research-driven policy and practice changes (*Figlio et al., 2015*).

UNITED STATES: EXAMPLES OF FAMILY-CENTERED SYSTEMS DESIGN

The following examples explore how integrated administrative data has been used for the purpose of enrolling children, families, and other residents in public programs such as for food, shelter, health, and child care.

Pennsylvania Integrated Eligibility System for Public Benefits

In 2003, Pennsylvania earned first place in the Center for Digital Government Best of the Web competition for innovations in State General Government for the state’s integrated online eligibility system, COMPASS (Commonwealth of Pennsylvania Access to Social Services). The online portal allows users to determine eligibility for Medicaid, Supplemental Nutrition Assistance Program, Temporary Assistance for Needy Families, Early Intervention Services, Subsidized Child Care, and many other public programs. When a resident enters his or her information into COMPASS, the data is integrated across programs, eliminating the need for information to be re-entered and re-verified (*Georgetown, 2008*).

The initial investment for development of COMPASS was approximately \$3.5 million, and it is maintained through cost allocation sharing split across four departments (Education, Health, Public Welfare, and Insurance.) COMPASS was initially conceived due to a 1990s lawsuit that resulted in bringing groups together to develop strategies for improving access to health care. Once success was demonstrated in health care efforts, more public agencies entered the COMPASS system and now there are dozens of programs that a resident can access through the portal.

In 2002, West Virginia became the first state to replicate the COMPASS system. They hired an independent organization to code transfer COMPASS into inROADS, West Virginia’s version of the application. It took a total of six months to design and develop inROADS. In addition, the following states have leveraged the COMPASS concepts, although execution varies among the states: Wisconsin, Delaware, Massachusetts, New Hampshire, Florida, Indiana, Georgia and New York (*Georgetown, 2008*).

Michigan Bridges

Michigan is the most recent state to create a family-centered integrated eligibility system using shared administrative data. The Michigan Bridges system was significantly updated in 2018, and is mobile friendly and available in English, Spanish, and Arabic. The new system allows for integrated applications for: child development and care, health, food, cash assistance, and emergency relief (such as housing). Bridges also allows families to take pictures of required eligibility documents with a mobile phone or tablet and instantly upload that document to the application. Additionally, Bridges serves as a resource for additional services and support and connects residents to hundreds of community-based organizations that provide support services. Community-based organizations have staff called “navigators” who are supported to help Michigan residents connect with needed resources in their home communities and to help them with the application if desired.

To better understand the successes and lessons learned of Michigan's Bridges eligibility system and resource platform, it is important to understand the original intent. In 2009, Bridges was launched as a self-service platform for food assistance. This was done, "... to help reduce the face-to-face intake process and divert these applications from case workers to regional processing centers. Case worker involvement was limited to final authorization... The primary business driver was high caseloads (600:1) and a statewide hiring freeze. These data points convinced policy leaders that a substantial broadening of this online service would yield even more significant results." (*Hogan, 2012*). In 2011, Bridges was expanded to all income-based eligibility systems and included more family-centric options, such as the uploading of documents and the inclusion of over 700 private organizations to provide application assistance for residents. A significant portion of the cost for the 2011 expansion was supported from private and philanthropic donations, including the active involvement of Council of Michigan Foundations.

Although the initial business driver of Bridges was staffing and cost efficiency for government, the use of integrated administrative data to consolidate eligibility determination has resulted in successful outcomes for its residents using public services. Some of these include: reduced time spent on applications for clients, high uptake of online applications, 90% customer report of preference for online channels versus face-to-face, reduced time of eligibility workers spent on administrative tasks, and increased numbers of eligible residents accessing services – which has resulted in increased economic benefit. (*Hogan, 2012*)

California's Enrollment Landscape

California does not have a family-centric eligibility system whereby administrative data is shared across various departments to take a family from application to enrollment. There are three main systems used in California, none of which actually complete the application process remotely, but some of which allow sharing of common data between departments. Of the 58 California counties, 39 counties use a system called "C4 Yourself," 18 counties use CalWIN, and Los Angeles uses a system called "Your Benefits Now." In all cases, these systems do not have the functionality of comparable systems from other states reviewed in this paper such as COMPASS, inROADS, and Bridges.

UNITED STATES: BIRTH MATCH REAL-TIME USE OF INTEGRATED ADMINISTRATIVE DATA

Although there are an extraordinary number of examples illustrating how cross-agency administrative data has been used for research, policy, and program planning, far fewer instances aim to make real-time, or near real-time, change that affects the health and well-being of young children. Birth Match is one of those cases. Birth Match aims to prevent maltreatment of newborns. Although operationalized differently in each of the three places that use Birth Match, the common purpose is to protect newborns from parents who have a history of neglect or abuse, including but not limited to having their parental rights terminated from children prior to the birth of the newborn. Michigan has the most comprehensive Birth Match process, followed by Maryland and New York City.

New York City

In April 2008, the New York City Administration of Children's Services experienced the death of an infant whose mother was already known by the foster system. Baby Pablo Paez, age 11 weeks, was tragically killed by his mother, who at the time had an open case in the New York child welfare system, as another child had been removed from her care by children's services. Pablo's mom was advancing toward reunification with her older child at the time of Pablo's death. In spite of the fact that child welfare staff were actively involved with her, no endeavors were made to inspect the well-being of the infant at or shortly after birth. Following Pablo's death, the Commissioner of the New York City child welfare organization issued a Child Safety Alert outlining the progression of incidents leading up to the late newborn child passing. The Alert was followed by a formal policy change at New York City Administration for Children's Services, whereby any child born to a parent who had a different child currently in foster care would lead to a child safety investigation. This was accomplished by comparing public health administrative data with social services administrative data (*Shaw, 2013*).

Maryland

Maryland's Birth Match enactment is more extensive in scope than New York City's and is statewide. The Maryland statute known as Birth Match (SB 421) was established by law on October 1, 2009. This law was initially conceptualized in 2004 when the Health Commissioner in Baltimore City, Peter Bielensohn, proposed the idea following the passing of twins by a mother who been known to the Department of Social Services (DSS) and who had had her parental rights ended on the grounds that she was judged by the courts as having inadequate ability to securely parent her children. Commissioner Bielensohn's introduction to the Birth Match idea did not have enough political sponsorship at the time, but the idea was resurrected in 2007 after another tragedy in Baltimore City. In 2007, a baby died from ingesting methadone while her mother was affected by drugs. The child's mother was already known to DSS; she had had two other children removed from her care by the foster system and her parental rights terminated.

In 2008, House Bill 1603 and Senate Bill 632 were proposed requiring the sharing of administrative information between the Maryland Department of Human Resources (Child Welfare) and the Department of Health and Mental Hygiene (Vital Statistics). This enactment was passed in the Senate, but then failed in the House due to privacy and legal concerns. At first, there was a belief that the birth records notification could be an encroachment on the privileges of parents in cases where they were known to child welfare services but there was insufficient proof of child maltreatment. There were also concerns about whether parents who voluntarily surrendered their parental rights would be incorporated into Birth Match.

In 2009, the enactment was reintroduced and passed. It required administrative information coordination between Child Welfare and Vital Statistics to alert Child Welfare of any instances of a parent who has a newborn and, within the last five years, has had their parental rights terminated. Each week, Child Welfare provides Vital Statistics with a list of parents whose rights have been terminated in the last five years. Vital Statistics then finds any matches with news births and sends notification to Child Welfare by the following day. When notified of new matches, Child Welfare sends a social worker to make contact with the family and assess newborn safety (*Shaw, 2013*).

Michigan

The administrative data birth coordination process in Michigan is both the most established and most thorough of the three places. Birth Match started in Michigan in 2001 between the Michigan Family Independence Agency, Department of Human Services (DHS) and the Michigan Department of Community Health Vital Records (DCH) to identify babies born to parents who have had an earlier termination of parental rights (dating back to 1978). Initially, DCH provided a listing of new births, including both the mother's and the father's names, to DHS each week.

In October 2008, Birth Match became state law (*Michigan HB 5814*). Michigan has computerized the procedure of administrative data coordination between DCH and DHS and has created comprehensive trainings for the child welfare workforce on the most proficient ways to investigate a Birth Match referral. While New York depends on case managers to recognize pregnancy or recent births and Maryland utilizes a weekly administrative data coordination process, Michigan operates on automated daily updates of births, which are analyzed for mothers and fathers with termination of parental rights and for adults with other cases of abuse, neglect, and other dangers to children such as sexual abuse. In Michigan, administrative data flows from Vital Records to Child Welfare daily, then Child Welfare cross references hospital records with parents with previously terminated rights or abuse/neglect history. In the case of a match, the state office automatically sends an email alert to the local office and an investigation is immediately required (*Shaw, 2013*).

California

California does not have Birth Match; however, it has analyzed administrative data for research and policy purposes and found directly relevant statistics. In a California cohort study of 1,576 newborns brought into foster care in 1995, the majority of them (nearly 1,000) had at least one sibling in foster care (*Shaw, 2013*). Data from across the United States shows that infants under the age of one are a disproportionate percentage of all children who die from abuse or neglect, 49.4% in 2015 (*Child Welfare Information Gateway, 2017*). Lastly, a presentation by Putnam-Hornstein in 2010 explored mortality among children reported for maltreatment and found that some types of prior child welfare involvement, for example physical abuse, dramatically increase the risk of intentional injury death. There was also a strong association between prior reports to child welfare and deaths of children under the age of five. In fact, these children are almost six times more likely to die than children without a previous child welfare report of intentional injury.

49.4%

Infants under the age of one are a disproportionate percentage of all children who die from abuse or neglect

GLOBAL EXAMPLES: ADMINISTRATIVE DATA AND INTEGRATED DATA SYSTEMS

Country-wide administrative data has the ability to produce a vast number of observations. It is possible to study population-level data, detect rare events that can be used for identification, and study the diverse impacts of educational practices and policies. Since the coverage is so wide, it is also possible to link administrative data from education to data from many other domains, such as health and workforce. This type of data also provides the opportunity to study other variables related to social networks, delinquency, geographical location, and health issues. Furthermore, studying intergenerational issues, such as comparing children's administrative records with those of their parents or grandparents is feasible when a country has a long history of data collection and integrated data analysis.

Nordic Countries

The study on the uses of administrative registers in Norway, Sweden, and Florida from the NBER Working Paper Series' Education Research and Administrative Data (*Figlio, Karbownik, & Salvanes, 2015*) produced interesting findings. Norway and Sweden have similar data structures; their administrative data are ideal for studying the long-term consequences of policy interventions.

In Norway, family links and parental background, combined with long panels of data, are used to study and analyze family and education policies, early investment, long-term outcomes in education and labor, and others. These findings paved the way for the government to focus on early investment in children, including parental leave policies.

In Sweden, there are four core data registers or sets: population (employment and education registers), activity (earnings or pupils registers), real estate (prices or GIS registers) and business (patent or schools registers). Aside from this data, researchers also use the unique individual identifiers of each Swedish resident. Because of this, the country is able to conduct empirical studies on family, education, labor markets, and health, as well as use data to inform the introduction of types of schools (e.g. charter schools) and policies (e.g. refugee placement policy).

Australia

Administrative data in Australia can be acquired through the Australian Bureau of Statistics (ABS), the Government's independent statistical agency. ABS maximizes the value of the data it collects by combining two or more sources together, also known as data integration. Information on Australian

families, communities, industry, and the economy are readily available, making it easier for the government to better understand the society and to deliver and improve its services.

The process of data integration is an efficient and effective way of combining data and creating new insights. Since ABS is an approved integrating authority, the agency can conduct data integration projects involving sensitive and confidential data. Various important case studies have already been conducted and published using data integration. The ABS website offers access to case studies on health, migrants, education, life expectancy, business, and others. ABS's Multi-Agency Data Integration Project offers case studies and research projects that combine information from multiple government agencies.

Canada

Administrative data is seen as necessary for policymakers in Canada to effectively understand their communities' needs and provide residents with better services. This led to the establishment of the Research Data Centres (RDC) Program, an initiative of Statistics Canada (the country's national statistical agency), the Social Sciences and Humanities Research Council (SSHRC), and the university consortia, to provide greater access to the data collected by Statistics Canada. The RDC network is funded by Statistics Canada, SSHRC, the Canadian Research Data Centre Network, Canadian Institutes of Health Research, and Canadian Foundation for Innovation. RDCs are located throughout the country.

RDCs give researchers and decision makers local, in-person access to confidential administrative and linked data, household surveys, and population microdata. There are over 100 data sets—from health surveys and child care surveys to schools and environmental surveys—that they can use for different purposes. This initiative is believed to help strengthen the country's social research and support the policy research community.

There are dozens of reports and findings available on the Canadian Research Data Centre Network, many of which link survey data to administrative data over time. For example, "Child Physical Punishment, Parenting, and School Readiness" (2018) utilizes administrative and survey data over a two-year period to link school readiness data, parenting practices, and physical punishment. A report produced in 2018 through a birth-census cohort called "Paternal Education and Adverse Birth Outcomes in Canada" concludes that low paternal education increases the risk of adverse birth outcomes, especially of fetal and infant mortality, independently from maternal characteristics.

Canada also provides an example of a regional effort to link and analyze administrative data specifically to improve programs and make better policy decisions for children. The Child and Youth Data Laboratory (CYDL) is a joint initiative between PolicyWise for Children & Families and government ministries in the province of Alberta that serve children. Through CYDL, researchers and policymakers can access different population data and analysis, such as Family Support for Children with Disabilities, Child Care Subsidy, Post-Secondary Education, Youth Corrections, and Income Support Transitions. Furthermore, CYDL, through its close partnership with the government, has conducted several longitudinal and special projects.

CONCLUSION

The advancement of integrated administrative data for action creates new opportunities to ask better questions that promote the right services and supports for families in a way that is more accessible to those families. It allows public funds to be targeted in a more efficient and effective manner. For governors and county-level decision makers, it promotes a clearer understanding of needs and assets, helping them to navigate competing resource demands. Legal and perceived obstacles, as well as a lack of leadership and capacity, have stymied many efforts to integrate data in the past. Nevertheless, there has been significant progress in the last five years, much of which has been initiated and supported by private philanthropy and then maintained by public systems. Examples from around the globe and from states, counties, and cities in the United States offer lessons and ideas for unlocking the potential of data to support equity among those with significant needs.

REFERENCES

- Allegheny County Department of Human Services. (2013). *Human Services Integration Fund*. Retrieved May 3, 2018, from <http://www.alleghenycounty.us/Human-Services/About/Funding-Sources.aspx>
- Carroli, G., Villar, J., Piaggio, G., Khan-Neelofur, D., Gülmezoglu, M. Mugford, M., et al. (2001). WHO systematic review of randomized controlled trials of routine antenatal care. *The Lancet*, 357(9268), 1565-1570.
- Center on the Developing Child at Harvard University. (2007). *A Science-Based Framework for Early Childhood Policy: Using Evidence to Improve Outcomes in Learning, Behavior, and Health for Vulnerable Children*. <http://www.developingchild.harvard.edu>
- Child Welfare Information Gateway. (2017). *Child abuse and neglect fatalities 2015: Statistics and interventions*. Washington, DC: U.S. Department of Health and Human Services, Children's Bureau.
- Commission on Evidence-Based Policy Making. (2017). *The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking*.
- Cox, M. Parent-child relationships. In M. Bornstein, L. Davidson, C. Keyes, and K. Moore (Eds.), *Well-being: positive development across the lifespan*. Hillsdale, NJ: Lawrence Erlbaum.
- Fraser, J. (2013). *Improving Educational and Well-Being Outcomes: School/DHS Data Sharing in Allegheny County, 2012 Update*.
- Figlio, D. N., Karbownik, K., & Salvanes, K. G. (2015). *Education Research and Administrative Data*.
- Figlio, D. N., Karbownik, K., Guryan, J., & Roth, J. (2014). *The Effects on Poor Neonatal Health on Children's Cognitive Development*.
- Georgetown University Policy Institute, Center for Children and Families. (2008). *Pennsylvania: Streamlined Enrollment & Renewal Through Technology*.
- Kitzmilller, Erika M. (2014). *IDS Case Study: Creating "One City": New York City's Center for Innovation through Data Intelligence*. Actionable Intelligence for Social Policy (AISP), University of Pennsylvania.
- LeBoeuf, W.; Barghaus, K.; Henderson, C.; Coe, K.; Fantuzzo, J.; and Moore, J. (2017). *The Use of Integrated Data to Inform Quality Pre-K Expansion in Philadelphia*. Research Briefs. 1. http://repository.upenn.edu/pennchild_briefs/1
- Lin, V., Maxwell, K., & Forry, N. (2017). *Considerations in Preparing to Analyze Administrative Data to Address Child Care and Early Education Research Questions*. OPRE Research Brief #2017-18. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- McGroder, S., Brooks, J., and LaPointe, P. (2016). *Improving Human Services Programs and Outcomes Through Shared Data*. Washington, D.C.: National Governors Association Center for Best Practices.
- National Research Council, Institute of Medicine. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Shonkoff, J.P. and Phillips, D.A. [eds.]. Washington, DC: National Academy Press.
- Reardon, S.F., Portilla, X.A. (2016). *Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry*. AERA Open. DOI: 10.1177/2332858416657343, Aug 2016.
- Rubin, D., O'Reilly, A., Zlotnik, S., Hendricks, T., Zorc, C., Matone, M., & Noonan, K., (2013). *Improving Education Outcomes for Children in Child Welfare*.
- Terry V. Shaw, Richard P. Barth, John Mattingly, David Ayer & Steve Berry. (2013). *Child Welfare Birth Match: Timely Use of Child Welfare Administrative Data to Protect Newborns*, *Journal of Public Child Welfare*, 7:2, 217-234.
- The Annie E. Casey Foundation. (2017). *Using Integrated Data Systems to Improve Case Management and Develop Predictive Modeling Tools*. Baltimore, MD: Author. Retrieved May 1, 2018, from <http://www.aecf.org/m/resourcedoc/aecf-usingIDStoimprovecasemanagement-2017.pdf>

This report was prepared by



VIVA Strategy + Communications is a social impact consulting firm providing transformative strategy and communications services to nonprofits, public agencies, and philanthropic organizations. Learn more at viva-strategy.com.

This report was created with support from



The Heising-Simons Foundation is a family foundation based in Los Altos, California. The Foundation works with its many partners to advance sustainable solutions in climate and clean energy, enable groundbreaking research in science, enhance the education of our youngest learners, and support human rights for all people. More at hsfoundation.org.

