

Advancing the Science of Children's Services through Large Data:

Convening Summary

March 2014

More than 100 people participated in a two-day convening held at the University of Southern California (USC) on February 27 and 28, 2014. The meeting served as a kick-off event for the Children's Data Network (CDN), a new data and research collaborative funded by First 5 LA, housed at USC's School of Social Work, and developed in collaboration with the California Child Welfare Indicators Project (CCWIP) at the University of California at Berkeley (UCB). Attendees representing the Los Angeles County Department of Children and Family Services, Department of Public Health, Office of Education, Public Information Office, Inter-Agency Council on Child Abuse and Neglect, Policy Roundtable for Child Care and Development, and Blue Ribbon Commission on Child Protection were joined by state-level colleagues from the California Department of Social Services, Department of Public Health, Health and Human Services Agency Office of Systems Integration, Administrative Office of the Courts, and Assembly Committee on Human Services. Also in attendance were more than a dozen foundation representatives with an interest in children's health and well-being, national and international researchers who traveled to participate in the discussion, and leaders from key LA-based partner organizations. Structured to provide a basis for further conversations with CDN stakeholders, sessions examined record-linkage centers from other countries, analytic methods employed with linked data, and the building blocks of successful university–agency–community data partnerships.

Day 1 opened with welcoming and introductory remarks from Dean Marilyn Flynn (USC School of Social Work), Armando Jimenez (director of research and evaluation at First 5 LA), and CDN codirectors Jacquelyn McCroskey (USC School of Social Work) and Emily Putnam-Hornstein (USC School of Social Work and CCWIP). Laying the groundwork for the presentations and conversations that followed, Dennis Culhane, professor of social policy at the University of Pennsylvania, distinguished between two key applications of integrated administrative data: its use for real-time service provision and its role in research and evaluation. Culhane presented rich case studies from both his own research and those compiled through the initiative he directs, Actionable Intelligence for Social Policy initiative, underscoring the potential for integrated administrative data to serve as the foundation for rigorous research leading to actionable and evidence-rich programs and policies.

Rebecca Glauert from the Telethon Institute for Child Health Research continued the discussion of integrated data by describing the Developmental Pathways Project, a long-standing, collaborative model for research and evaluation in Western Australia. Emphasizing the importance of partnerships and community engagement

Children's Data Network

coupled with scientific rigor, Glauert's presentation covered issues ranging from the architecture and governance of integrated data to the ultimate dissemination of findings. She highlighted the importance of using data linkage as an opportunity to break down government silos, promote informed policy decision-making, and generate new knowledge.

Noting that "bigger" data do not necessarily translate into greater knowledge, Fred Wulczyn, senior research fellow at Chapin Hall at the University of Chicago, provided candid reflections on the importance, and challenge, of attaching relevant meaning to data and research. He noted that absent meaning and understanding, numbers are just numbers. Offering a cautionary tale of examinations that fail to incorporate structural factors as determinants of individual and agency behaviors, Wulczyn stressed the importance of understanding individual-level dynamics in context and the value of system-centric analysis.

In the final session of Day 1, the audience reviewed concrete examples of research generated through integrated administrative data. Melissa Jonson-Reid, professor and director of the Brown Center for Violence and Injury Prevention at Washington University, presented her research examining children's cross-system involvement based on 15 administrative data sources. Findings indicated that many children identified by schools as needing special education services were seen, on average, 5 to 8 years earlier by child protective services (CPS). Jonson-Reid leveraged this knowledge to work with community partners from early childhood, child welfare, and special education to improve protocols for connecting children and families referred to CPS for in-home parenting support services, helping to bridge communication gaps between agencies.

Peter Fallesen, a researcher at the Rockwool Foundation Research Unit in Denmark, provided a second international perspective on integrated data concerning children and families. Noting his initial confusion by the very concept of "linked data" in the United States, he explained that it is known much more simply in Denmark as "data," because all records related to individuals are integrated using a universal identifier by default. Fallesen continued to describe findings based on the rich data available for studying children and service systems.

Mark Courtney, professor at the School of Social Service Administration at the University of Chicago, broadened the data linkage conversation by discussing research based on administrative data linked with survey and other ad hoc data sources. Courtney outlined methodological opportunities for using administrative sources to cost-effectively standardize data elements and track survey participants over time, providing examples of various studies that have utilized this approach. In his presentation, he reflected on the strengths of administrative data, but also noted that marrying multiple administrative sources can be affected by poor matching, missing data due to population mobility, and poor data entry practices.

Children's Data Network

Day 2 turned to a somewhat controversial area—predictive risk modeling using information harvested from administrative data. Moira Wilson, principal analyst of New Zealand's Ministry of Social Development, provided an overview of the country's current exploration of predictive risk modeling as a tool for allocating place-based resources and strategically targeting children at high risk of maltreatment. Rhema Vaithianathan, a professor of economics at New Zealand's Auckland University of Technology, discussed the use of predictive risk modeling in health care settings for such purposes as assessing patient risk of re-hospitalization following discharge. Advantages include the cost-effective nature of the process, the absence of a human factor to sway outcomes, the placement of risk scores along a continuum, and the assessment of risk based on population-level data. Vaithianathan noted, however, the tremendous challenges of implementation and that the success of risk modeling is dependent upon not only the information that can be used for modeling purposes, but also effective and available preventive interventions.

Offering a different perspective, Irene De Haan, a frontline practitioner and adjunct professor from New Zealand's University of Auckland, discussed the ethical issues associated with predictive risk modeling in the context of child maltreatment. She emphasized the importance of delivering targeted services without destabilizing or stigmatizing families. Wilson, Vaithianathan, and De Haan provided a rich perspective from multiple vantage points—government worker, academic researcher, and social work practitioner.

David Schwartz, director of research and evaluation at the William Penn Foundation, explored the potential to use integrated administrative data to improve upon some of the existing risk-assessment tools currently used in children's services. He presented findings based on "machine-learning" algorithms to build predictive risk models that performed in a superior fashion to traditional regression approaches. John Fluke, associate director for research at the Kempe Center, reflected on decision-making contexts and the potential role of predictive models. He discussed how integrated data may aid in the prevention of child maltreatment, but was careful to point out that there will always be false positives (cases in which intervention occurred but was not needed) and false negatives (cases in which intervention was needed but did not occur). Fluke raised questions that are essential to address when considering predictive risk modeling: "How common is the outcome?" and "What trade-offs exist if the wrong answer is acted upon?" He also tested the audience's threshold for mistakes by posing two hypotheticals of false positives and false negatives so the consequences could be weighed.

Following a dynamic conversation regarding predictive risk modeling, Denise Herz, professor and director of the School of Criminal Justice and Criminalistics at California State University, Los Angeles, returned to the theme of deriving meaning from data. Detailing the lives of families and children who create the stories behind administrative data, Herz reminded convening participants over lunch that crunching numbers from

Children's Data Network

administrative data help uncover the “who, what, where, and when,” but other sources are often needed to delve into the “why” and to make meaning of the data.

The convening concluded with a diverse panel of speakers that reflected on the data and research landscape, described critical areas in children's programming and policy in which knowledge is needed, and presented candid assessments of the challenges and opportunities facing partners in the CDN. The panel was introduced by Barbara Needell, principal investigator of CCWIP and chair of the California Child Welfare Council's Data Linkage and Information Sharing Committee, who emphasized the importance of the work proposed by the CDN. John Kim, codirector of the Advancement Project, underscored the importance of collaboration and ensuring the incorporation of perspectives and voices from the community into research questions and priorities. Arturo Delgado, superintendent of the L.A. County Office of Education, provided an educational perspective on the importance of timely access to cross-systems information meaningful to school principals, teachers, and school support staff. Michael Wald, emeritus professor of law at Stanford University, issued a provocative challenge for the CDN to identify essential outcomes that all children should have achieved as they transition to adulthood, as well as a small number of core indicators that could be consistently tracked over time. Shell Culp, director of the Office of Systems Integration for California's Health and Human Services Agency, shared the state's innovative approach to promoting interoperability and open data among state departments. Finally, Greg Rose, deputy director of the California Department of Social Services, Children and Family Services Division, presented some of the challenges of data integration within state-level government and between state and county agencies, while making a strong case for the necessity of integrating information in California.

Throughout both days, themes and common insights emerged. During each presentation, speakers underscored the critical difference between using linked records for research and evaluation (i.e., the goal of the CDN) and potential government uses of linked records for real-time service delivery and decision-making concerning individual clients. A second theme related to the importance of strong collaboration among university, agency, foundation, and community partners to ensure the relevance of research and its ultimate translation into policies and programs. Finally, both speakers and participants exhibited genuine enthusiasm and interest regarding the creation of the CDN. As one person reflected in a follow-up email: “This is a project that is much needed and very exciting—please know that I will help in any way I can.” This sentiment was echoed by many others and will undoubtedly prove to be a critical element for the Children's Data Network.