

Children's Data Network

Supporting a Strong Start for CA Kids:

A Proposal to Augment, Connect, Support, and Collect Population-Based Data

The Children's Data Network (CDN) is a data and research collaborative focused on the linkage and analysis of administrative records. With faculty, research scientists, and students spanning multiple universities and an intensely collaborative approach to partnership with the public sector,¹ we are able to draw upon a wealth of both substantive and methodological expertise. Much of our work is prospective; starting from birth we use integrated data to follow children forward to characterize their experiences, public service interactions, and outcomes. We also act as an extension of state and local government, helping to envision, build, and sustain the systems that make the underlying linked data available for research and analysis.²⁻¹⁰ Ultimately, our goal is to make better use of data to inform and improve programs and policies.

Our work at the nexus of the academic and public sectors and reputation as trusted partners in this space¹¹⁻¹³ offers a unique vantage point from which to view existing and emerging analytic, operational, and policy opportunities related to young children and their families. From where we stand, we see the great data work happening in this space, and, as a result, know that much can – and has yet to – be said about young children and their families. We also have come to understand that not all critical questions about this important population can be asked and answered with the data in its current (or future) form.

In the brief scope of work that follows, we outline our plan to attend to these gaps in understanding of young children in California. By augmenting, connecting, promoting, and collecting population-based integrated data, we can better ensure all young children in California are born with a strong start.

Background. Working in partnership with colleagues in the public sector, both as key advisors on cross-sector data integration efforts and as trusted analysts generating data about young children, has prompted a number of take-aways. Notably:

Much can be said about the health, safety, and public service interactions of young children involved with public benefit and service systems in California. About seven years ago, California Health and Human Services (CalHHS), an agency that administers essential services to millions of our state's most vulnerable residents, began to formally transition from a program- to client-centric model of service delivery.¹⁴ For that transition to be successful, however, they needed client-centered data. We are proud to have helped them fill that gap by conducting the first and ongoing agency-wide record reconciliations that link and organize administrative, client-level records across major CalHHS programs and data sources (i.e., CalFresh (aka Supplemental Nutrition Assistance Program); CalWORKs (aka TANF); Child Protection; Developmental Services; Emergency, Hospital Discharge, and Ambulatory Surgery; Family Planning, Access, Care, and Treatment (Family PACT); In-Home Supportive Services; Medi-Cal; Vital Birth; Vital Death; and Women, Infants, and Children (WIC)). This data integration effort has resulted in the development of encrypted master "intra-agency" client identifiers, which facilitates the exchange of statistical information both within and between individual CHHS departments and the generation of important information about service involvement at a client and population level.^{2,5,15} Now in its 5th year, the data submission process has been normalized, the production of cross-program data has been routinized, and the value of the person-level, longitudinal datasets for collaborative research, evaluation, and everyday operations has been made clear. We believe that the success of this data integration initiative and the current state of the resulting data now affords exciting opportunities to perform more targeted analyses – analyses that characterize the health, safety, and public service interactions of young children, specifically.

Currently, less can be said about their experiences and engagement in early care and education and other family support services. A number of critically important efforts, however, are currently under development

Children's Data Network

that aim to characterize the early care and education of children ages 0-5 and to collect, in the future, individual-level data.¹⁶⁻²⁰ Each new data-related and technological development moves us closer to a more robust, child-centric understanding early childhood in California. A current and expansive analysis of the status and trajectory of these efforts, how they work together, the type of data collected, and key partners would be instrumental to identifying prospects for future integration with existing child-level data and opportunities for collaboration, guidance, and support.

The existing integrated administrative data tell us that, despite the resources available in California, a substantial proportion of young children are at risk for poor outcomes. Our research has highlighted the incredible variation in the likelihood that children will experience adversities, including maltreatment, unintentional and intentional injuries, and death before age 5.²¹⁻²⁵ We also have piloted a number of data-based approaches to identifying those children most at risk for these objectively poor outcomes so that they can be more easily prioritized for enhanced services and supports.^{6-8,26-34} In this process, and owing to a population-based frame, we have also found that, at times, the very programs designed to help and support young children are not reaching the families who need them most.^{35,36}

Administrative data, however, will always be limited in its ability to give us information from the perspective of the children and families served. First 5 California, County Commissions, other stakeholders, and local service providers have invested significant resources in developing, marketing, administering, and evaluating services for maximum benefit. Given the program-focused nature of administrative data collection and the collection of information via survey methods that do not allow for individual-level linkage,^{37,38} however, there is currently no population based or family-centered view of the service landscape across California communities. There are no data that speaks to the barriers, benefits, and challenges new parents may face in accessing the supports they need. There is no risk-adjusted understanding of who chooses to engage in services and characterization of who we're missing. In order to be able to address those gaps and ensure that all children have a strong start in life, we need to hear from all families about their needs, barriers, and experiences.

Opportunity. Given existing integrated data, we can say a lot about young children and their families. That data, however, is mainly limited to those who interact with public benefit and service systems. Our lack of understanding about the needs and experiences of young children and families *who don't engage, but at are at high risk of poor outcomes* hobbles efforts to align needs and supports, improve programs, and refine engagement strategies. Ultimately, these limitations constrain not only our ability to say, at a high-level, how young children in California are faring, but also hinder our ability to effectively evaluate programs and strategically connect and amplify the impact of our investments. Effectively, the administrative data in its current (and future) form will always be incomplete, limiting our ability to gauge impact and address inequities that could propel strategic investments by First 5 California for years to come.

Proposal. During the next 3 years (August 1, 2023 – July 31, 2026) and at an anticipated cost of \$3M, we propose to attend to our gaps in understanding of young children in California by augmenting, connecting, supporting, and collecting population-based integrated data.

Approach. Our work can be divided into four main workstreams that we will pursue simultaneously:

Stream 1. Augment our Understanding of Young Children using Available Integrated Data: Use the data that is available, including the Strong Start Index, to better characterize this priority, but understudied, population.

Stream 2. Connect Ongoing Data Initiatives through a Data-Focused Landscape Analysis: Describe the data that are and are not currently available in data systems related to the entire population of young children

Children's Data Network

ages 0-5 in California, and, to the extent possible, assess implications for potential data integration and collaboration.

Stream 3. Support through Targeted Technical Assistance around Record Linkage and Analysis: Support and guide the development of data and data systems critical to furthering that child-level and population-based understanding, including the CalHHS Record Reconciliations, Agency Data Hub, Brilliant Beginnings, Cradle-to-Career, the Whole Child Equity Tool, and others.

Stream 4. Collect information about Service Needs and Experiences Directly from Families: Collect critical information concerning the service needs and experiences of families with young children from parents who recently gave birth in California. Specifically, we will build upon the administrative data profile of children and families in California that was developed through the Strong Start Index to develop a representative sample of children born into families with the greatest needs. With approval from the CalHHS Committee for the Protection of Human Subjects and the Vital Statistics Advisory Board, we will use information from the Strong Start Index and associated birth records to contact a select set of new parents with the goal of enrolling them in direct data collection efforts. Information from parents will be gathered at several time points during the first two years of the child's life to not only better understand the landscape of potential family needs and the reach of our current services, but also to give voice to the experiences of the families of young children. With parental and agency consent, we also plan to, to the extent possible, augment self-reported family experiences with records of public benefit and ECE utilization, including records in CalHHS Research Data Hub, Brilliant Beginnings, and the Cradle-to-Career Data System.

Timeline. We have outlined the initial three-year workplan, but anticipate that opportunities for continued engagement with families for additional data collection over time will emerge.

Year 1: (August 1, 2023 – July 31, 2024)

Stream 1. Augment our Understanding of Young Children using Available Integrated Data

Planned Activities:

- Updating Strong Start Index website to include 2022 records and additional actionable geography (birthing hospital)
- Working in collaboration with First 5 Commission to explore enhanced/equity-targeted Strong Start Visualizations

Potential Deliverables: (Categorized by audience type)

- Updated Strong Start Index website with additional geographies and reconceptualized equity-targeted dashboard (*External*)
- Memo describing new findings, features, and use cases (*Internal*)
- Updated Technical Documentation (*Technical*)

Stream 2. Connect Ongoing Data Initiatives through a Data-Focused Landscape Analysis

Planned Activities:

- Initiating early childhood data-focused landscape analysis

Potential Deliverables: (Categorized by audience type)

- Infographic of Key Findings related to the Landscape of State-Level Early Childhood Data (*External*)
- Report on the Landscape of State-Level Early Childhood Data (*Internal*)
- Detailed Appendix with program information (*Technical*)

Stream 3. Support through Targeted Technical Assistance around Record Linkage and Analysis

Planned Activities:

- Conducting CalHHS RR to integrate 2022 records

Children's Data Network

- Exploring ways to highlight young children in data visualizations
- Providing targeted TA to, among others, California Health and Human Services Center for Data Insights and Innovation (CDII) to take on the responsibility of annual record reconciliation and analysis of cross-program records and to build and maintain the Agency Data Hub;^{1,15} Cradle-to-Career (C2C) Data Governing Board (via Data and Tools Advisory Board¹¹); Brilliant Beginnings; the Whole Child Equity Tool;^{19,20} and other emerging early-childhood related data initiatives

Potential Deliverables: (Categorized by audience type)

- Cross-Program Dashboard posted in CalHHS Open Data Portal (*External*)
- Early Childhood Data Dashboard submitted to CalHHS and First 5 CA for review (*Internal*)
- Technical Documentation (*Technical*)

Stream 4. Collect information about Service Needs and Experiences Directly from Families

Planned Activities:

- Developing a population-based sample based on the Strong Start Index
- Seeking data approvals
- Engaging subawardees to facilitate analysis
- Partnering with First 5 CA and Commissions to promote data collection effort

Potential Deliverables: (Categorized by audience type)

- Study communication materials (*External*)
- Project update memo (*Internal*)
- Human Subjects Protocol (*Technical*)

Year 2: (August 1, 2024 – July 31, 2025)

Stream 1. Augment our Understanding of Young Children using Available Integrated Data

Planned Activities:

- Updating Strong Start Index data and visualizations

Stream 2. Connect Ongoing Data Initiatives through a Data-Focused Landscape Analysis

Planned Activities:

- Finalizing early childhood data focused landscape analysis

Stream 3. Support through Targeted Technical Assistance around Record Linkage and Analysis

Planned Activities:

- Facilitating CalHHS Record Reconciliation and Early Childhood Data Dashboard; Ongoing targeted TA

Stream 4. Collect information about Service Needs and Experiences Directly from Families

Planned Activities:

- Piloting data collection, finalizing outreach protocol, launching of formal data collection

Year 3: (August 1, 2025 – July 31, 2026)

Stream 1. Augment our Understanding of Young Children using Available Integrated Data

Planned Activities:

- Updating Strong Start Index data and visualizations

Stream 2. Connect Ongoing Data Initiatives through a Data-Focused Landscape Analysis

Planned Activities:

Children's Data Network

- Socializing and publishing early childhood data focused landscape analysis

Stream 3. Support through Targeted Technical Assistance around Record Linkage and Analysis

Planned Activities:

- Facilitating CalHHS Record Reconciliation and Early Childhood Data Dashboard; Ongoing targeted TA

Stream 4. Collect information about Service Needs and Experiences Directly from Families

Planned Activities:

- Collecting data, ongoing analysis, and synthesis of emerging findings

Children's Data Network

References

1. Foust R, Hoonhout J, Eastman A, et al. The Children's Data Network: Harnessing the Scientific Potential of Linked, Administrative Data to Inform Children's Programs and Policies. *International Journal of Population Data Science*. 2021;6(3). doi:10.23889/ijpds.v6i3.1702
2. Putnam-Hornstein E, Ghaly M, Wilkening M. Integrating Data To Advance Research, Operations, And Client-Centered Services In California. *Health Affairs*. 2020;39(4):655-661. doi:10.1377/hlthaff.2019.01752
3. Children's Data Network. CHHS Annual Record Reconciliation. Accessed September 30, 2021. <https://www.datanetwork.org/news/chhs-annual-record-reconciliation/>
4. Children's Data Network. *California Health and Human Services Agency (CalHHS) DRAFT Record Reconciliation Methodology*; 2021. https://www.datanetwork.org/wp-content/uploads/RR-2020_Basic_Methodology.pdf
5. California Health and Human Services (CHHS) Open Data Portal. CalHHS Record Reconciliation Cross-Program Dashboard Data (2015-2020). Accessed September 24, 2021. <https://data.chhs.ca.gov/dataset/chhs-record-reconciliation-dashboard-data-2015-2020>
6. Los Angeles County Risk Stratified Supervision Model – Children's Data Network. Accessed September 28, 2021. <https://www.datanetwork.org/research/los-angeles-county-risk-stratified-supervision-model/>
7. Children's Data Network. Assessing Children's Risk Using Administrative Records: A Proof of Concept Predictive Risk Modeling (PRM) Project – Children's Data Network. Accessed September 30, 2021. <https://www.datanetwork.org/research/assessing-childrens-risk-using-administrative-records-a-proof-of-concept-predictive-risk-modeling-prm-project/>
8. Cuccaro-Alamin S, Foust R, Vaithianathan R, Putnam-Hornstein E. Risk assessment and decision making in child protective services: Predictive risk modeling in context. *Children and Youth Services Review*. 2017;79:291-298. doi:10.1016/j.childyouth.2017.06.027
9. Children's Data Network. *Predictive Risk Modeling: Findings from California's Proof-of-Concept*. https://www.datanetwork.org/wp-content/uploads/PRM_CWDAB_2018.pdf
10. Bill Text - AB-2083 Foster youth: trauma-informed system of care. Accessed November 18, 2021. https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB2083
11. California S of. Data & Tools Advisory Board. C2C. Accessed January 24, 2023. <https://c2c.ca.gov/data-tools-advisory-board/>
12. cdnadmin. Key Children's Data Network Staff Tapped to Advise the California Cradle-to-Career Data System – Children's Data Network. Accessed January 24, 2023. <https://www.datanetwork.org/news/key-childrens-data-network-staff-tapped-to-advise-the-california-cradle-to-career-data-system/>
13. Children's Data Network. CDN Work/Partnership Showcased at CHHS 2021 Data Expo. Accessed September 30, 2021. <https://www.datanetwork.org/news/cdn-work-partnership-showcased-at-chhs-2021-data-expo/>
14. Guiding Principles and Strategic Priorities. California Health and Human Services. Accessed February 5, 2023. <https://www.chhs.ca.gov/guiding-principles-strategic-priorities/>
15. Children's Data Network. CalHHS Record Reconciliation and Research Data Hub. Accessed September 24, 2021. <https://www.datanetwork.org/research/chhs-record-reconciliation-and-research-data-hub/>
16. *California Brilliant Beginnings Data Landscape*; 2021. https://cappa.memberclicks.net/assets/PublicPolicy/2021/California%20Brilliant%20Beginnings%20Data%20Investment_5.18.21.pdf

Children's Data Network

17. California Department of Social Services. *Child Care Transition Quarterly Report*; 2023. https://www.cdss.ca.gov/Portals/9/CCDD/Jan%202023%20Quarterly%20Report%20Final_1_4.pdf?ver=2023-01-04-163841-897
18. State of California Cradle-to-Career System. Cradle-to-Career System Public Data Definitions. C2C. Accessed January 19, 2023. <https://c2c.ca.gov/public-data-definitions/>
19. Gov. Newsom signed AB 2832 – here's what that means for California families | Catalyst California. Accessed April 12, 2023. <https://www.catalystcalifornia.org/blog/gov-newsom-signed-ab-2832--heres-what-that-means-for-california-families>
20. Bill Text - AB-2832 Whole Child Community Equity. Accessed April 12, 2023. https://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=202120220AB2832
21. Children's Data Network. Population-based, Birth Cohort Studies of Child Deaths Before Age 5. Accessed September 30, 2021. <https://www.datanetwork.org/research/population-based-birth-cohort-studies-of-child-deaths-before-age-5/>
22. Putnam-Hornstein E. Report of Maltreatment as a Risk Factor for Injury Death: A Prospective Birth Cohort Study. *Child Maltreat*. 2011;16(3):163-174. doi:10.1177/1077559511411179
23. Putnam-Hornstein E, Cleves MA, Licht R, Needell B. Risk of Fatal Injury in Young Children Following Abuse Allegations: Evidence From a Prospective, Population-Based Study. *Am J Public Health*. 2013;103(10):e39-e44. doi:10.2105/AJPH.2013.301516
24. Putnam-Hornstein E, Schneiderman JU, Cleves MA, Magruder J, Krous HF. A Prospective Study of Sudden Unexpected Infant Death after Reported Maltreatment. *The Journal of Pediatrics*. 2014;164(1):142-148. doi:10.1016/j.jpeds.2013.08.073
25. Children's Data Network. Cumulative Risk of Child Protective Service Involvement before Age 5: A Population-Based Examination. Accessed September 30, 2021. <https://www.datanetwork.org/research/cumulative-risk-of-child-protective-service-involvement-before-age-5-a-population-based-examination/>
26. Vaithianathan R, Rouland B, Putnam-Hornstein E. Injury and Mortality Among Children Identified as at High Risk of Maltreatment. *Pediatrics*. 2018;141(2). doi:10.1542/peds.2017-2882
27. Children's Data Network. *Preliminary Risk Validation Findings*; 2020. <https://www.datanetwork.org/wp-content/uploads/preliminary-risk-validation-findings.pdf>
28. Putnam-Hornstein E, Foust R, McCroskey J, Dalton E, Vaithianathan R. Predictive Risk Modeling: Practical Considerations. :2.
29. Children's Data Network. Leveraging Hospital Records to Examine Child Maltreatment Risk. Accessed September 30, 2021. <https://www.datanetwork.org/research/leveraging-hospital-records-to-examine-child-maltreatment-risk/>
30. Children's Data Network. Developing the California Strong Start Index to Inform and Monitor Program Investments and Policy Development for Children and Families in California. Accessed September 28, 2021. <https://www.datanetwork.org/research/developing-the-california-strong-start-index-to-inform-and-monitor-program-investments-and-policy-development-for-children-and-families-in-california/>
31. Foust R, Prindle J, Eastman AL, et al. Using Universal Vital Records to Monitor Children's Assets at Birth: The Strong Start Index. Published online 2022. <https://www.datanetwork.org/wp-content/uploads/Strong-Start-Index-Methodology-Streamlined-8.3.22.pdf>
32. Strong Start Index. Accessed September 30, 2021. <https://strongstartindex.org/>

Children's Data Network

33. View Key Findings – Strong Start Index. Accessed December 12, 2022. <https://strongstartindex.org/view-key-findings>
34. Children's Data Network. Helping Los Angeles County Chart a Course Toward Universal and Targeted Home Visiting. Accessed September 30, 2021. <https://www.datanetwork.org/research/helping-los-angeles-county-chart-a-course-toward-universal-and-targeted-home-visiting/>
35. Foust R, Prindle J, Eastman AL, et al. *Antelope Valley Partners for Health Home Visiting Linkage Memo.*; 2019. <https://www.datanetwork.org/wp-content/uploads/HV-Memo-Revised-Sept-2019.pdf>
36. Putnam-Hornstein E, Prindle J, Hammond I. Engaging Families in Voluntary Prevention Services to Reduce Future Child Abuse and Neglect: a Randomized Controlled Trial. *Prev Sci.* 2021;22(7):856-865. doi:10.1007/s11121-021-01285-w
37. California Health Interview Survey | UCLA Center for Health Policy Research. Accessed February 5, 2023. <https://ask.chis.ucla.edu/chis/Pages/default.aspx>
38. Maternal and Infant Health Assessment. Accessed February 5, 2023. <https://www.cdph.ca.gov/Programs/CFH/DMCAH/MIHA/Pages/default.aspx>