What Can We Learn About the Incidence of Foster Care Placement from Birth Records?

Findings from the Cross Jurisdiction Model Replication Project

September 2022

OPRE Report 2022-255

Authors:
Emily Putnam-Hornstein, University of North Carolina at Chapel Hill, Children's Data Network; Eunhye Ahn, Washington University in St. Louis, Children's Data Network; and Rhema Vaithianathan, Auckland University of Technology, Children's Data Network

Jared Parrish and Robyn Husa, Alaska Division of Public Health

Matthew Walton and James P. Perry, Kentucky Office of Data Analytics

Claire Smither Wulsin, M.C. Bradley, and Beth Varley, Mathematica

I. Introduction

Accurate and ongoing public health surveillance of the incidence of child maltreatment and related risk and protective factors can help programs and policymakers as they work to shape prevention and intervention efforts. One approach to capturing this information is to link local, state, and/or federal administrative records, such as those from child welfare, health, social services, education, public safety, and other agencies. The Child Abuse Prevention and Treatment Act (CAPTA) requires the examination of a variety of topics related to the incidence of child abuse and neglect with the aim of better protecting children from maltreatment and improving the well-being of maltreatment victims. The Child Maltreatment Incidence Data Linkages (CMI Data Linkages) project consists of two components that explore how innovative administrative data linkages can improve our understanding of the prevalence of child maltreatment and identify related risk and protective factors. The first component was a study to assess the feasibility of using linked administrative data to improve understanding of child maltreatment and related risk and protective factors. The feasibility study identified promising practices and contextual and organizational factors that may promote the use of linked administrative data. This brief describes the Cross Jurisdiction Model Replication (CJMR) project, the second component of the CMI Data Linkages project.

About risk prediction

Risk prediction modeling is an analysis method that uses existing data to predict the likelihood of outcomes. Risk prediction models are a promising tool to guide decision-making about services and care at the individual level and to guide the allocation of resources at the state, county, or community level. For example, in the context of child protection, agencies have used risk prediction models to inform child welfare agency primary prevention efforts (Daley et al. 2016). Agencies also have used these models to understand geographic differences in rates of recorded maltreatment, given variations in populations and detection systems (Drake et al. 2020). However, some risk prediction models have faced challenges related to transparency, bias, and data availability that limit their usefulness.◢

The CJMR project sought to understand the degree to which a risk prediction model built from population-level and anonymized birth records in one state could be used to differentiate the risk of foster care placement in other jurisdictions.[1] Specifically, the CJMR project estimated population-level differences in the risk of being placed in foster care by applying a single risk prediction model (from California) to anonymized birth records from Alaska and Kentucky. The risk prediction model was designed to explore whether birth record data could help jurisdictions to understand geographic variation in risk, improve planning for services, and focus limited resources on the most at-risk communities.

Specifically, the CJMR project sought to address the following questions:

- Can information recorded at birth consistently predict the risk of foster care placement?
- Can a model developed (trained) using data from one jurisdiction be used by other jurisdictions to estimate the share of children who might have a heightened likelihood of foster care placement?

[1] See a discussion of the development and use of risk prediction models in child welfare in Allegheny County, Pennsylvania, at https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx

- What opportunities exist for using birth record models to help inform the ongoing and accurate national surveillance of foster care placement? That is, can these models be used to understand differences in foster care placement across counties and states?

The CJMR project examined the potential for state child protection agencies to characterize geographic differences in foster care placement risk using existing birth record data, with the goal of informing efforts to plan and focus community-based resources. The CJMR project also examined whether a risk prediction model implemented in multiple states (tailored, as needed) could improve the accuracy of ongoing national surveillance of child maltreatment.

Some risk prediction models implemented by state child protection agencies have faced challenges around transparency, bias, data availability, and data quality; this is particularly true for models intended to guide decision making on an individual level. To mitigate these challenges, the CJMR model was built in collaboration with child protection agencies, using anonymized data that many agencies already have on hand (that is, birth record data), to address population-level epidemiological questions while upholding community engagement and transparency at every step (e.g., the code is open source on GitHub).

This project was conducted as a collaboration between Mathematica, the Children's Data Network (CDN) (consisting of researchers at the University of North Carolina at Chapel Hill; University of California, Berkeley; Auckland University of Technology; and Washington University in St. Louis), the Alaska Division of Public Health, and the Kentucky Office of Data Analytics.

The CJMR risk prediction model is hosted on a GitHub site (https://github.com/childrensdatanetwork/PRM-birth-cohort). Other jurisdictions are invited to use or adapt the model and add their work to this site.

II. Overview of Participating Jurisdictions

This section provides relevant context on each participating jurisdiction. The technical appendix provides more details about each jurisdiction and other information to supplement this overview.

Exhibit 1. Overview of participating jursdictions

	California	Alaska	Kentucky
2020 number of child victims with substantiated or indicated maltreatment	60,317	3,212	16,748
2020 rate of child victims with substantiated or indicated maltreatment per 1,000 children	6.9	18	16.7
Level of governance	County administered	County administered	State administered
State agency	California Department of Social Services	Alaska Department of Family and Community Services	Kentucky Cabinet for Health and Family Services

Source: Children's Bureau 2021, 2022; Child Welfare Information Gateway 2018.

Note: A child victim is a child for whom the state determined at least one maltreatment was substantiated or indicated, and a disposition of substantiated or indicated was assigned for a child in a report. This includes a child who died, and the death was confirmed to be the result of child abuse and neglect. A child might be a victim in one report and a nonvictim in another report.

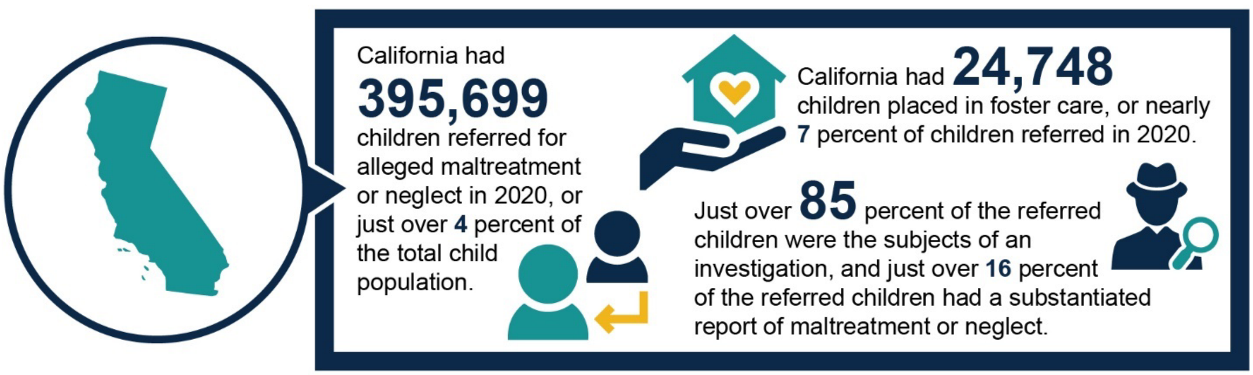
A. California

1. How is the child protection system organized?

California uses a county-administered model of governance. The California Department of Social Services (CDSS) is the agency with overall responsibility for the state's child welfare system. Each of the state's 58 counties administers its own child protection system. CDSS performs regulatory oversight, develops policies and regulations, provides training and support, and monitors compliance and performance through an outcomes accountability system (California-Child and Family Services Review).

County child welfare agencies oversee cases from investigation to permanency. This includes foster care services or screening, assessments, and services to keep the child safe in the home. The agencies support family reunification or develop permanency plans, including adoption and guardianship placements, if children cannot be safely returned to their biological families (County Welfare Directors Association of California n.d.).

Exhibit 2. Snapshot of current child protection statistics in California



Source: Children's Bureau 2021, 2022.

See technical appendix for more details.

How does the child protection system use administrative data?

California aims to be a national leader in data-driven policymaking, data transparency, and public access to data (for example, see the State Assembly Bill AB636). California has codified these goals in legislation, and CDSS has formed several strategic partnerships for data sharing and analysis.

What is the analytic capacity of child protection agency staff?

The California Budget Act of 2011 shifted tax revenue from the state to counties for funding child protection, foster care, adoption, and adult protection. This realignment resulted in increased responsibility for CDSS to monitor outcomes for these service populations. Since then, CDSS has reorganized several times to meet this greater need for accountability. Today, the Research, Automation & Data Division works to provide accurate and timely data to inform policies for child protection and other social service programs serving vulnerable Californians. The Center for Data Insights and

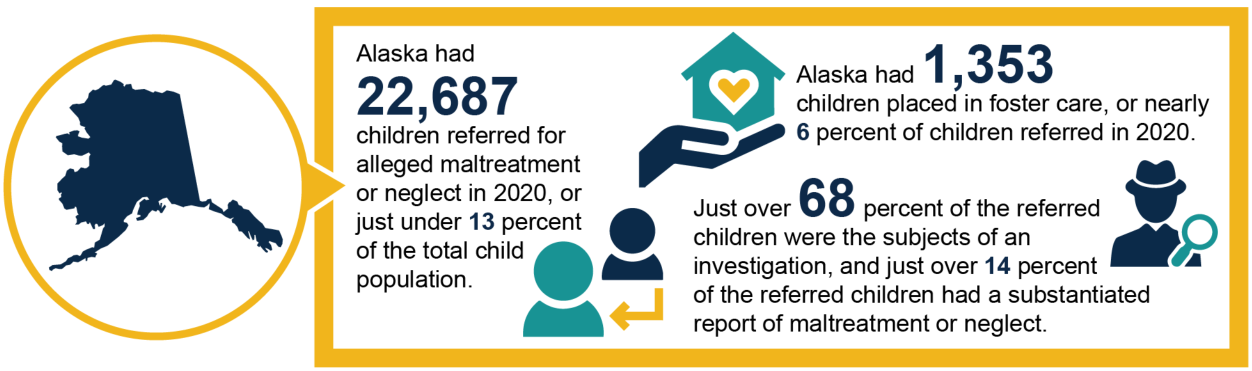
Innovation, which is part of CDSS's parent agency, the California Health & Human Services Agency, also supports CDSS in helping to link data across state service systems. The California Child Welfare Indicators Project and CDN partnerships provide analytic services for CDSS, including linking and analyzing administrative data. Both partnerships include staff with expertise in various statistical programs and processes.

B. Alaska

1. How is the child protection system organized?

Alaska's child protection system is run by the Office of Children's Services (OCS) in the Alaska Department of Health and Social Services, which was split into two departments on July 1, 2022. The OCS is now located in the Alaska Department of Family and Community Services. There are 21 regional offices divided between four service regions.

Exhibit 3. Snapshot of current child protection statistics in Alaska



Source: Children's Bureau 2021, 2022.

See technical appendix for more details.

How does the child protection agency use administrative data?

OCS is a data-driven agency and regularly uses administrative data for practice interventions, to develop policies and procedures, and to engage internal and external stakeholders. In addition, OCS uses its data system to track costs of services in relation to special needs costs, augmented rates, and other categories for requests for funds, foster care payments, and subsidies. OCS uses administrative data to generate random case samples and to extract population cohorts. This data enables staff to more thoroughly understand the issues bringing families to the attention of OCS, the reasons for substantiated maltreatment, permanency trends, categories of abuse and neglect for intakes, and investigations. OCS partners with Alaska's Division of Public Health (DPH) to integrate child welfare system administrative data with epidemiological and other statewide administrative data sets, such as birth and death records, Medicaid data, and education data. OCS staff use these linked data to improve internal processes and better understand populations that are likely to become involved with the child welfare system.

What is the analytic capacity of child protection agency staff?

The agency uses data, in conjunction with child welfare best practices, to guide all work. OCS has a research unit that produces data reports for staff to use. These reports are organized by staff level (line worker, administrator, supervisor, or manager) and then by module. The research unit also pulls ad hoc data as requested. The Online Resource for the Children of Alaska (ORCA) is built on the SQL Server, and therefore all data elements within ORCA can be queried. OCS also has a Continuous Quality Assurance team that analyzes statistics and creates reports for leaders and staff.

To expand the scope of analytics to the entire Alaska population, Alaska DPH partnered with OCS to create the Alaska Longitudinal Child Abuse and Neglect Linkage project (ALCANLink). The Alaska DPH Maternal Child Health Epidemiology Unit links and analyzes epidemiological and administrative data to address a variety of questions relevant to public health prevention efforts and child welfare activities.

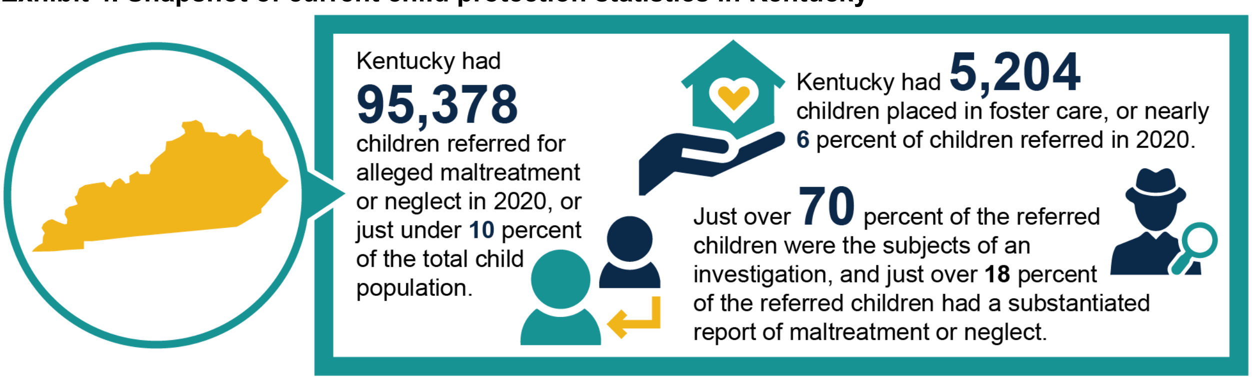
C. Kentucky

1. How is the child protection system organized?

Kentucky's child protection system is a state-administered program under the Department for Community Based Services (DCBS) within the Kentucky Cabinet for Health and Family Services (CHFS). DCBS administers the state's adult protective services system, as well as Kentucky's child protective services (CPS) programming, Kentucky's foster care system, and relationships with contracted service providers (such as those providing mental health services for families involved with CPS). DCBS's role in Kentucky has also recently expanded to assisting Supporting Kentucky's Youth, a Medicaid managed care contract program that consolidates and tailors services for youth in foster care.

There are 120 counties in Kentucky; each receives CPS services from a local county office. Those 120 counties are divided into nine service regions (see graphic here). A service region administrator leads each service region and supports the county-level supervisors and DCBS child welfare staff who report to them.

Exhibit 4. Snapshot of current child protection statistics in Kentucky



Source: Children's Bureau 2021, 2022.

See technical appendix for more details.

How does the child protection system use administrative data?

Kentucky uses administrative data for research and evaluation of child welfare intervention programs and operational enhancements. For example, one project involves linking child welfare data to the Kentucky Health Information Exchange portal to enable CPS workers to access health information for children in out-of-home care (Walton et al. 2022). There are also operational linkages between The Worker Information System and education data, Medicaid administrative claims data, and other state administrative data systems to enable the work of DCBS as it serves children in out-of-home care (Hall et al. 2021).

What is the analytic capacity of child protection agency staff?
The CHFS secretary has prioritized the secure and efficient sharing of data for the purposes of evaluation and quality improvement. CHFS leaders have also placed special emphasis on using data analytics to diminish health disparities in Kentucky. CHFS contracts with local universities to hire technical professionals who have training in data governance, analytics, statistics, research methods, database architecture, and information security.

Kentucky's analytic capacity is distributed across multiple agencies. Each constituent department of CHFS has some staff devoted to data management and analysis. However, the Office of Data Analytics (ODA) and the Office of Application Technology Services are two of the primary centers for database management, research, and data analytics for projects about state health and social services. ODA supports CHFS operations with technical analysis, program evaluation, research, and policy analysis. The goal is for CHFS to be a resource for the practice of evidence-based policymaking. ODA analysts link data and collaborate with members of the university and research communities to promote health and social services research.

III. Data Sources

To be considered for the CJMR project, jurisdictions were required to have preexisting access to child protection data and birth records. However, each jurisdiction's records did not need to focus on the same time period. Exhibit 5 shows the specific data sources each jurisdiction used.

Alaska and Kentucky were selected as replication jurisdictions based on the following characteristics:

- Access to birth records and child protection data
- Staff with the technical skills needed to link the data and implement the risk prediction model
- Time, resources, and interest in participating

Exhibit 5. Data sources by state

Name	Years	Geography covered	Source	Measures
California				
Child protection data	2012–2018 (development) 2000–2003 (validation)	California (statewide)	California Department of Social Services, Child Welfare Services Case Management System	Foster care placement
Birth records	2012–2015 (development) 2000 (validation)	California (statewide)	California Department of Public Health	Birth characteristics used to develop model predictors
Alaska				
Child protection data	2006–2022[a]	Alaska (statewide)	Alaska Department of Health and Social Services	Foster care placement

Name	Years	Geography covered	Source	Measures
Birth records	2013–2016	Alaska (statewide)	Alaska Department of Health and Social Services, Division of Public Health	Birth characteristics
Kentucky				
Child protection data	2015–2021[b]	Kentucky (statewide)	Department for Community Based Services, The Worker Information System	Foster care placement
Birth records	2015–2021	Kentucky (statewide)	Kentucky Cabinet for Health and Family Service, Department for Public Health, Office of Vital Statistics	Birth characteristics

[a] Alaska Department of Public Health did not create an age restriction to allow agencies to include additional fetal maltreatment cases (that is, reports recorded while the mother was still pregnant with the child in our cohort, with the unborn child as the victim).

[b] Kentucky Office of Data Analytics selected these years to ensure that at least five years of outcomes data were available for children born in 2015, and at least three years of outcome data were available for birth cohort years 2015–2017.

IV. Methods

The risk prediction model used birth records to explore whether a uniform set of characteristics could consistently differentiate risk of foster care placement. The model was not designed to identify individual children who would experience foster care placement. Instead, it aimed to use prediction methodologies to understand population-based risk differences in foster care placement.

The CJMR project took place in three stages: development, validation, and replication (Exhibit 6). CDN developed a model using recent birth record data from California. This model was then validated using two data sets unrelated to the project sample: (1) records for children born in California more than two decades ago (that is, children born in 2000) and (2) records for children born in Allegheny County, Pennsylvania. To examine the generalizability of the model and the ease with which the statistical program could be used to code a common set of birth record predictor variables, Alaska and Kentucky replicated the model using data from their jursidictions.

Exhibit 6. Stages of the work: Development, validation, and replication

A. Development

In the development stage, the CDN team developed a model trained to classify differences in risk of foster care placement. CDN constructed the risk prediction model using population-wide birth records in California from 2012–2015 linked at the individual level to administrative child protection records from 2012–2018. The target outcome for model training purposes was a foster care placement before the child's third birthday. The cumulative incidence of a first foster care placement by age 3 for these birth cohorts was 1.8 percent. The three-year follow-up window for developing predictions enabled CDN to create a model based on children who were born recently (between 2012–2015) and who had foster care placements that could be consistently observed via linked records. This follow-up period also enabled CDN to examine the year of life when maltreatment reporting, substantiation, and foster care placement rates are highest (that is, during infancy) and captured the period when about 70 percent of maltreatment fatalities occur (that is, before age 3) (Child Welfare Information Gateway 2021).

To develop the model, CDN used 75 percent of the records as a training set, and kept the remaining 25 percent of records for testing the model. Thus, 1,490,917 records were randomly selected for the training process, and 496,972 were used to test the model. The model was built with 82 coded predictor variables derived from fields captured in birth records. Those predictors reflect what was known about the characteristics of the child, the parents, past pregnancies, current conditions, and other factors at the time of birth. A full list of predictor variables is available in the technical appendix.

B. Validation

The model, trained using California records from recent birth cohorts, was validated using two data sets unrelated to the original cohort. One data set of children born in California in 2000 was used to examine whether the model could generalize to birth cohorts in the same jurisdiction but from different time periods. A second data set of children born in Allegheny County, Pennsylvania from 2012–2015 was used to examine whether the model could generalize to birth cohorts from the same period but different jurisdictions. Across both data sets, the model's ability to classify differences in a child's likelihood of future foster care placement was weaker than with the training data but still sufficient (AUC > 0.750; see Section V for explaination of AUC).

1. Validation using the California 2000 birth cohort

CDN created the same predictor variables using birth records from 2000 by following the steps taken to make the variables for the model. Because the previous predictor variables were coded using data from the 2012–2015 birth records, some modifications were needed for the 2000 records. This included a crosswalk of data elements that had been given different names or whose coding had changed over time in the birth data (such as the categories for maternal education). In addition, some predictor variables could not be generated because the data elements were not available in 2000 birth records (such as maternal smoking).

After coding all predictor variables in the birth records, CDN used linked child protection records to construct the same outcome used for modeling purposes: a binary variable showing whether or not there was a foster care placement within three years of birth.

2. Validation using the Allegheny County, Pennsylvania 2012–2015 birth cohort

Individual birth records included all children who were born from 2012 to 2015 in Alleghany County and each child's foster care placement status through age 3. The model was applied to the Allegheny County cohort and proved more accurate at differentiating risk than a similar model created using Allegheny's own smaller data set (see the Allegheny County website for additional details: https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx). This may be because the more data that is available to train a risk prediction model, the more likely it is to be representative and generalizable to other populations.

C. Replication

The final stage of the CMJR project involved applying the model code to two other jurisdictions: Alaska and Kentucky. The goals of this stage were twofold: (1) to see if other jurisdictions could implement the model code and apply it to their own anonymized birth data with limited technical assistance and (2) to understand whether the model performed similarly in different jurisdictions with distinct child welfare contexts. Alaska and Kentucky implemented the model using programming code from CDN. Each jurisdiction generated the predicted probability of future foster care placement and then looked to see how many children who were classified at different risk levels had documented foster care placements in their linked child protection records.

The Alaska and Kentucky child protection and birth records data have the same general structure and organization as the California data. Therefore, during the replication stage, Alaska and Kentucky did not change the model other than to reflect a few differences in available data. For example, because states determine what is on their birth record forms, all states might not collect the same data (National Research Council 2009). Alaska and Kentucky also did not attempt to train a new or different model. If the California birth record included a predictor variable not included in Alaska or Kentucky birth record data, the team dropped that variable from the code. See the technical appendix for a full list of variables used by each jurisdiction.

V. Findings

Findings highlight the potential for leveraging administrative data to understand geographic differences in foster care placement risk. This work may inform efforts to plan and focus community-based resources and help identify populations experiencing disparities across jurisdictions. Findings also suggest that a risk prediction model could be used to understand the varying incidence of foster care placement across jurisdictions, as measured through administrative child protection records. Further, a model that is implemented in multiple states (even if slight tailoring is needed) could provide a population-based foundation for understanding the relationship between various risk factors at birth and the associated risk of foster care placement.

A. Feasibility of replicating the model in Alaska and Kentucky

Alaska and Kentucky ran the model using their existing child protection and birth record data without significant modification. CDN provided minimal technical support to troubleshoot code errors.

The Alaska team recoded some variables and revised some categorization from variables to align with the model. For example, the maternal and parental education variables in Alaska were missing one of the categories from the California variable. The Alaska team created additional categories so that the variables would be similarly structured.

The Kentucky team also recoded some variables to align with the format of variables in the model. For example, race, payment, and fetal birth presentation variables all required reassignment of categorical values to match the original model. Other Kentucky variables offered more categorical values than California. For example, Kentucky birth record data includes six sub-designations for one of several types of Medicaid payments; all of these were aggregated into a single variable.

B. Model performance results

The teams evaluated model performance using area under the curve (AUC) and recall score. AUC is a measure of the ability of a model to distinguish between a binary outcome (that is, an outcome with only two possible values), where 1 indicates perfect accuracy and .5 indicates total randomness. Generally, the higher the AUC, the better the model performs at predicting the binary outcome. A score between 0.7 and 0.8 is considered an acceptable AUC score, between 0.8 and 0.9 is an excellent score, and greater than 0.9 is an outstanding score (Mandrekar 2010). Using California data for births between 2012 and 2015 (the data on which the model was built), the model has an AUC of 0.870.

Exhibit 7. Recall scores by risk decile rank and AUC by state

	California 2012–2015		Alaska 2013–2016		Kentucky 2015–2022	
	(AUC = 0.870)		(AUC = 0.801)		(AUC = 0.790)	
Decile rank	Risk score threshold	Recall score	Risk score threshold	Recall score	Risk score threshold	Recall score
10	0.68	59.4%	0.71	34.6%	0.81	32.8%
9	0.54	75.7%	0.57	57.6%	0.69	56.8%
8	0.44	84.8%	0.48	72.6%	0.56	73.4%
7	0.35	90.9%	0.41	84.2%	0.48	83.9%
6	0.26	95.4%	0.32	92.1%	0.40	90.9%
5	0.16	98.2%	0.23	97.4%	0.32	95.8%
4	0.09	99.1%	0.15	99.3%	0.23	98.5%
3	0.04	99.7%	0.08	99.8%	0.12	99.8%
2	0.02	99.9%	0.03	100.0%	0.05	100.0%
1	0.00	100.0%	0.00	100.0%	0.00	100.0%

AUC = area under the curve.

The two replication sites (Alaska and Kentucky) had similar AUCs. The 2013–2016 Alaska birth cohort had an AUC score of 0.801 (Exhibit 7). The AUC score was slightly lower using the 2015–2022 Kentucky birth cohort, at 0.790.

Recall scores, also known as the true positive rates, indicate how many children were accurately predicted to be placed in foster care among all children who entered foster care.

Exhibit 7 shows the risk scores and recall scores by deciles. When all children in the birth cohort are ranked by their predicted risk of being placed in foster care, the ten percent with the lowest risk scores are in decile one, at the bottom of the table. Using with the California 2012–2015 cohort, these children have a predicted risk of 0.00 percent (risk score threshold) and the model correctly predicted their foster care placement 100 percent of the time (recall score). The next lowest decile have predicted risk scores between 0.00 – 0.02. The model correctly predicted their foster care placement 99.9 percent of the time. Jumping to the top ten percent of children with the highest risk score, these children have a predicted risk of foster care placement of 0.68 or greater. Of all the children placed in foster care who are in the top ten percent of predicted risk, the model correctly identified 59.4 percent of children placed in foster care. The model did not correctly predict the remaining 40.6 percent of foster care placements for children in the top ten percent of predicted risk.

For the Alaska 2013-2016 cohort, Exhibit 7 shows that children in the top ten percent of predicted risk had a predicted risk of 0.71 or greater (risk score threshold) and that the model correctly predicted their foster care placement 34.6 percent of the time (recall score). The second ten percent of most at-risk children had a predicted risk between 0.57 – 0.71 and the model correctly predicted their foster care placement 57.6 percent of the time. Of all the children placed in foster care who are in the top ten percent of predicted risk, the model correctly identified 34.6 percent of children placed in foster care. The model did not correctly predict the remaining 65.4 percent of foster care placements for children in the top ten percent of predicted risk.

For the Kentucky 2015-2022 cohort, Exhibit 7 shows that children in the top ten percent of predicted risk had a predicted risk of 0.81 or greater (risk score threshold) and that the model correctly predicted their foster care placement 32.8 percent of the time (recall score). The second ten percent of most at-risk children had a predicted risk between 0.69 – 0.81 and the model correctly predicted their foster care placement 56.8 percent of the time. Of all the children placed in foster care who are in the top ten percent of predicted risk, the model correctly identified 32.8 percent of children placed in foster care. The model did not correctly predict the remaining 67.2 percent of foster care placements for children in the top ten percent of predicted risk.

C. Implications for the field, possible next steps, and limitations

Research Question 1. Can information recorded at birth consistently predict risk of foster care placement?

Findings suggest that birth data can be used to model and differentiate risk of foster care placement. The CJMR project found that the risk prediction model created from 2012–2015 California birth record data was able to accurately predict foster care placement within the first three years of life.◢

Research Question 2. Can a model developed (trained) using data from one jurisdiction be used by other jurisdictions to estimate the share of children who might have a heightened likelihood of foster care placement?

The risk prediction model created from California birth record data generalized well to two other jurisdictions: Kentucky and Alaska. These jurisdictions had different levels of governance and different structures for their child welfare agencies from the jurisdiction used to develop the model.◢

Research Question 3. What opportunities exist for using birth record models to help inform the ongoing and accurate national surveillance of child maltreatment? That is, can these models be used to understand differences in foster care placement across counties and states?

With this model, jurisdictions could use only birth records (even if that data is not linked to anything else) to predict foster care placement. This risk prediction model might be useful for jurisdictions that do not have the analytic capacity or data accessibility to develop their own models. Not all jurisdictions have the same analytic capacity as California, Alaska, and Kentucky to link individual birth records and child protection records to generate a data set and replicate the model. ◢

This risk prediction model may be a useful source of information to help guide prevention and intervention efforts and to inform new and existing policies. For example, the model may identify populations or jurisdictions where additional prevention investments may be warranted by revealing communities with the greatest need. It also may inform the planning and implementation of service delivery across diverse geographies and communties. Ultimately, implementation of the risk prediction model may improve the accurate and ongoing surveillance of the incidence of foster care placement. Additional jurisdictions with the analytic capacity to create a linked data set of individual birth records and child protection records may consider implementing the model with their own data to continue testing.

Finally, additional work is needed to better understand how well the model performs relative to models generated by individual jurisdictions; this was beyond the scope of the current project. Notably, this model does not account for the role of each state's child protection definitions and policies. States can vary widely in terms of how they define maltreatment, the size of their workforce, the capacity of their foster care system, and the practices of their court systems, among other policies. This variation may affect how the model performs within and its utility to particular jurisdictions.

References

Children's Bureau. "Child Maltreatment 2020." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau, 2022. https://www.acf.hhs.gov/sites/default/files/documents/cb/cm2020.pdf.

Children's Bureau. "Child Welfare Outcomes Report Data: Entered Foster Care During FY." Adoption and Foster Care Analysis and Reporting System (AFCARS). 2021. https://cwoutcomes.acf.hhs.gov/cwodatasite/enteredCare/index.

Child Welfare Information Gateway. "Child abuse and neglect fatalities 2019: Statistics and interventions." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau, 2021. https://www.childwelfare.gov/pubpdfs/fatality.pdf.

Child Welfare Information Gateway. "State vs. County Administration of Child Welfare Services." Washington, DC: U.S. Department of Health and Human Services, Children's Bureau, 2018. https://www.childwelfare.gov/pubPDFs/services.pdf.

County Welfare Directors Association of California "Children's Services." n.d. https://www.cwda.org/childrens-services-0

Daley, Dyann, Michael Bachmann, Brittany A. Bachmann, Christian Pedigo, Minh-Thuy Bui, and Jamye Coffman. "Risk Terrain Modeling Predicts Child Maltreatment." *Child Abuse & Neglect,* vol. 62, 2016, pp. 29–38. https://doi.org/10.1016/j.chiabu.2016.09.014.

Drake, B., M. Jonson-Reid, M.G. Ocampo, M. Morrison, and D. Dvalishvili. "A Practical Framework for Considering the Use of Predictive Risk Modeling in Child Welfare." *The ANNALS of the American Academy of Political and Social Science,* vol. 692, no. 1, 2020, pp. 162–181. https://doi.org/10.1177/0002716220978200.

Hall, M.T., M.T. Walton, R.A. Huebner, G.E. Higgins, A.B. Kelmel, and D. Lorenz. "Sobriety Treatment and Recovery Teams for Families with Co-Occurring Substance Use and Child Maltreatment: A Propensity Score-Matched Evaluation." *Children and Youth Services Review,* vol. 131, 2021.

Mandrekar, J. N. (2010). Receiver Operating Characteristic Curve in Diagnostic Test Assessment. Journal of Thoracic Oncology, 5(9), 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d.

National Research Council. "Vital Statistics: Summary of a Workshop." Michael J. Siri and Daniel L. Cork, rapporteurs. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press, 2009.

Walton, M.T., K.L. Conner, and J.A. Mackie. "The Future of Child Welfare Will Look Toward Children's Health: The Case of the Kentucky Health Information Data Sharing (KHIDS) Project." *Journal of Public Child Welfare,* vol. 16, no. 4, 2022, pp. 549–554.

September 2022

OPRE Report #2022-255

Project Officers: Christine Fortunato and Jenessa Malin

Project Specialist: Hanna Katz
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Project Director: Matt Stagner
Mathematica
111 East Wacker Drive, Suite 3000
Chicago, IL 60601-4303

Suggested citation: Emily Putnam-Hornstein, Eunhye Ahn, Rhema Vaithianathan, Jared Parrish, Robyn Husa, Matthew Walton, Claire Smither Wulsin, M. C. Bradley, and Beth Varley (2022). What Can We Learn About the Incidence of Foster Care Placement from Birth Records? Findings from the Cross Jurisdiction Model Replication Project. OPRE Report #2022-255, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services

This brief was funded by the Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, under contract number HHSP233201500035I/HHSP23337022T