U.S. Department of Health and
Human Services,
Office of the Assistant Secretary for
Planning and Evaluation

Predictive Analytics in Child Welfare

An Assessment of Current Efforts, Challenges and Opportunities

October 2017

By Christopher Teixeira and Matthew Boyas

The MITRE Corporation

This report was prepared by staff of the MITRE Corporation as operator of the CMS Alliance to Modernize Healthcare Federally Funded Research and Development Center under a project sponsored by the Assistant Secretary for Planning and Evaluation. The findings and conclusions of this report are those of the authors and do not necessarily represent the views of ASPE or HHS.

Executive Summary

Child welfare agencies are interested in leveraging new and emerging techniques to help them harness data and technology to make dramatic improvements to child welfare practice and ultimately produce better outcomes for children and families. This document explores the state of the use of predictive analytics in child welfare by conducting an environmental scan of child welfare agencies, academia, nonprofit organizations, and for-profit vendors. Topics discussed in qualitative interviews included how each jurisdiction uses predictive analytics to support child welfare practice, the challenges that motivated using predictive analytics, and the challenges faced as these agencies have begun their modeling efforts. Whether motivated by an unfortunate event or a directive to improve performance on a certain metric, agencies recognize that multiple problems may need to be tackled to obtain the full benefit of predictive analytics.

Interviews reveal that currently child welfare agencies most often use predictive analytics to address four key problem areas:

1. *Estimating elevated risk during or following preventive services.* These efforts seek to estimate future high risk of maltreatment, serious injury, or a child fatality.
2. *Forecasting the likelihood of repeated events.* This issue area looks at repeat maltreatment of children by perpetrators where the unit of analysis could either be the child or perpetrator.
3. *Analyzing cross-system interactions.* These efforts describe problems that exist with people or entities outside of child welfare and problems that often require support from outside agencies. This includes looking at children who are abused that might abuse their own children as a young parent or evaluating a neighborhood's risk level for abuse.
4. *Providing insight into agency operations.* This final category includes efforts to help organizations use predictive analytics to look at their own activities, such as evaluating caseworker turnover or trends in the quantity of incoming cases.

Significant challenges remain to using predictive analytics in the child welfare field. Predictive analytics are intended to supplement good casework practice and administration by providing data-driven insights into relative risks and the interactions among diverse circumstances. While predictive analytics can be the backbone of a system to systematically identify high-risk cases, no agency staff we spoke with believes that decisions should be automated. Agencies generally understand that predictive analytics is an additional tool to add incrementally to the child welfare toolkit. However, significant challenges still include, for example, ethical concerns related to the idea that predictive models, because of their reliance on historical patterns, may be inadvertently biased against certain subgroups of the population. To mitigate potential ethical concerns and to avoid potential biases, child welfare agencies prefer to keep ethical concerns in mind when choosing problems to model, in addition to being careful about the datasets included in the models and how they are used.

Data are the most crucial part of a predictive analytics implementation, and all agencies cited data acquisition and quality as the most challenging aspect of the implementation. Agencies run into two main challenges surrounding data: 1) obstacles to data sharing, both statutory and the sometimes over-cautious interpretations of those statutes, and 2) inconsistencies around data quality.

After the agency has acquired the requisite data, the next step in an implementation of predictive analytics is to choose whether to conduct the analyses in-house or through vendor contracts. Agencies recognize that they do not currently have in-house staffing capacity to develop or maintain applications on their own, but contracting with a third party nonprofit/for-profit provider, whether via a short-term or long-term contract, comes with other concerns around lack of control. Regardless of the methods of model development, predictive analytics is a costly venture with high upfront expenses, and agencies expressed worry about developing long-term funding strategies to provide the necessary support for the efforts.

Overall, the implementation of predictive analytics in child welfare is in its infancy across child welfare agencies. Some agencies plan on allowing caseworkers access to detailed risk scores, while others plan on restricting the use of risk scores for screening purposes in the agency's call center. No agencies were willing to comment on the accuracy and impact of their predictive analytics initiatives, citing that not enough time has passed since implementation to be confident in such metrics.

Significant training is also necessary to help caseworkers effectively understand the predictive modeling techniques and the interpretation of risk scores. In addition, agencies have expressed concern that without adequate service options to ameliorate risk, risk scores could lead to risk aversion and foster care caseloads could rise. In the long term, this could affect whether sufficient resources will be available to maintain predictive analytics systems.

During the interviews, child welfare agencies identified areas in which the federal government may be able to reduce some of the common challenges agencies face when implementing predictive analytics. These include:

- Developing templates for necessary legal documents, such as data sharing agreements between local-level agencies.
- Providing documentation and technical assistance to agencies as they develop contracts for hiring outside predictive analytics vendors.
- Specifying guidance on the additional data elements, and common data definitions, that could be added to those currently being captured in the Statewide Automated Child Welfare Information System (SACWIS) / Comprehensive Child Welfare Information System (CCWIS) data to make systems more suited to predictive analytics applications.
- Preparing sample predictive analytic scripts or programs for common problems.
- Establishing a secure central repository, hosted by a non-governmental trusted third party, for child welfare agencies across the country to upload and analyze data.
- Supporting agencies' efforts to build internal capabilities, either financially or through learning collaboratives or other in-kind developmental opportunities.

Interviews indicated that some agencies may have unrealistic expectations about the near term value and impact of predictive analytics. Predictive analytics alone will not provide a panacea for all challenges child welfare agencies face but can be an important tool to support processes, practices, policies and other systemic changes and improvements. For example, most of the child welfare agencies interviewed reported some degree of improvement in conjunction with implementing their predictive analytics solutions. However, they also indicated there will be an extended learning process before agencies mature in their understanding of the strengths and limitations of predictive models and how to best incorporate them into the other tools, strategies and interventions needed to improve child welfare outcomes. Other industries have struggled through similar challenges in adapting predictive analytics solutions to their processes, but as their methods and understanding of predictive analytics matured, entities were able to see significant benefits for the use of predictive analytics in accomplishing their mission. While the child welfare industry's application of predictive analytics remains in its infancy, predictive analytics in child welfare can mature similar to other industries through the sharing of experiences, best practices and evaluation.

Table of Contents

Introduction

This document assesses the current state and use of predictive analytics in the child welfare field. It was developed by staff of The MITRE Corporation for the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, through the CMS Alliance to Modernize Healthcare (CAMH) Federally Funded Research and Development Center (FFRDC). MITRE staff performed an environmental scan examining how predictive analytics is being used in the field of child welfare. Given the advancements in the field of predictive analytics, there is a strong interest in its potential to enable child welfare practitioners to make more informed decisions to ensure children are able to grow and thrive in a nurturing home environment, avoid tragic outcomes such as a serious injury or fatality due to abuse and neglect, and allocate resources strategically in a constrained economic situation. This document explores initial efforts underway in a number of communities pioneering these techniques and highlights the promise and challenges experienced by agencies undertaking these projects. Agencies and researchers working in this area discussed many challenges including technical limitations with poor data quality, ethical barriers in using certain data elements, and political obstacles in sharing data between local-level agencies. This document concludes by identifying potential opportunities where the Federal Government may be able to support child welfare agencies in their adaptation of predictive analytics solutions.

Methods

To develop the best approach for characterizing the state of predictive analytics in child welfare practice, CAMH worked with partners in both ASPE and the Administration for Children and Families (ACF) to identify five child welfare agencies that are recognized for using or evaluating the use of predictive analytics in their practice. In addition, CAMH reached out to four vendors to understand how they partner with child welfare agencies, as well as any barriers or limitations to those partnerships. Project staff also conducted a literature search to identify peer reviewed publications on predictive analytics in child welfare. Based on this literature search, CAMH identified and contacted three researchers to learn more on their work with child welfare agencies. These activities identified sixteen entities engaged in various activities related to predictive analytics in child welfare. Of these entities, CAMH successfully contacted fourteen of them, and conducted full interviews with twelve. The full list of interviewed entities can be found in the Discussion schedule.

While this list of agencies, vendors, and researchers comprises a significant breadth of the predictive analytics applications in child welfare, this is not a complete list, nor did everyone approached provide information. Project staff sought to eliminate bias by choosing the interviewees to maximize the diversity of the sample in terms of the size of the agency/group, level-of-involvement with predictive analytics, and geographic location within the United States. Based on the conversations, CAMH does not anticipate that those that did not respond to the request for information would have provided contradictory information to what is presented in this document. Instead, additional conversations likely would have provided additional nuance and details to the assessment.

Definition of terms

In the context of this document, predictive analytics is defined as analysis that uses data and algorithms to answer the question "Given past behavior, what is likely to happen in the future?" A form of advanced analytics, predictive analytics often employs sophisticated statistical techniques to use information about the past and present to predict what is likely to occur at some time in the future.[1] While some child welfare agencies are currently involved in what can be best described as descriptive or diagnostic analytics projects which are focused on providing data-driven snapshots of the child welfare system as it currently exists, other agencies are increasingly turning towards predictive analytics to help implement preventive interventions before a problem escalates. However, not all agencies want to independently develop and maintain this predictive capability in house.[2] Many agencies partner with nonprofits, independent experts, or academic researchers and/or for-profit companies to accomplish their goals.

It is important to mention that while this document is specifically focused on *predictive* analytics, child welfare agencies are simultaneously engaging in a range of advanced analytics projects. Many are increasingly incorporating data into their decision making processes, including using data to inform and refine traditional actuarial-based risk scoring tools. Agencies are also embracing the field of data visualization by creating dashboards and visuals of their data for executives, enabling quick access to reporting. These differ from predictive analytics in that dashboards often provide information about the current status of an agency or a particular child's case, enabling reactionary behavior; predictive analytics, on the other hand, can provide a forward-looking view into what might occur in the future, encouraging preventive behavior before that probabilistic future becomes reality.

Problem assessment

The specific problems child welfare agencies seek to address through predictive analytics vary widely. Based on the interviews conducted with the diverse collection of respondents described above, the problems that child welfare agencies have chosen to address with predictive analytics solutions fit into four primary categories:

1. Estimating *elevated risk* of maltreatment, serious injury, or a child fatality during or following preventive services.
2. Forecasting the likelihood of *repeated events,* such as maltreatment of children or re-entry into the foster care system.
3. Analyzing *system issues,* including cross-system interactions between the child welfare system and individuals or external agencies.
4. Providing insight into *agency operations,* such as evaluating caseworker turnover or trends of incoming cases.

[1] Definition of predictive analytics adapted from Gartner, Inc. (2016). Gartner glossary
[2] The Analytics Maturity Model (2015). ibm.com

These four categories can be further broken out into nine more specific problems or challenges, each one with a slightly different problem description and focus of analysis. These specific challenges, along with the overarching categories to which they map, are outlined below in

Table 1. A Detailed list of problems being modeled appears in the Appendix at the end of this document and contains more information on the problems, the data used, and current project status of predictive analytics applications in child welfare that the project team identified.

Table 1. Problem categories addressed in child welfare applications of predictive analytics

Problem category	Problem description	Unit of analysis
Elevated Risk (during or following preventive services)	Risk of chronic/severe maltreatment, serious injury, or fatality	Child and perpetrator (unique pair)
Repeated Events	Repeat maltreatment of children by perpetrators already in the system Re-entry of a child into the foster care system	Child or perpetrator (singular) Child
System Issues	Abused children becoming perpetrators Unfortunate outcomes after achieving permanency (e.g., homelessness, incarcerations) Demographic and crime-based predictions of regionalized risk of abuse	Child Child Geographic Regions
Agency Operations	Worker turnover Caseload size What/When/How much service is needed to achieve a better outcome	Caseworker Caseworker Child

The following sections delve into each of these four categories in more detail, providing insights into why agencies chose to focus on these problems. Concerns or limitations identified by child welfare agencies are addressed in the section below headed "Ethical considerations and other constraints."

Elevated risk during or following preventive services

Interviews combined with a review of the literature and media accounts suggest that the problem category "elevated risk" is the most common class of problems motivating agencies' explorations of predictive analytics modeling. This category is somewhat broad; it covers all instances of elevated risk involving children and/or perpetrators during or following preventive services. Many jurisdictions interpret this as 'elevated risk for a fatality' but the likelihood of a child experiencing foster care placement or the likelihood of significant maltreatment/near-fatality is also an outcome that would fall into this category and may employ a similar analytical approach. More often than not, agencies typically decide to address a problem in this category as a direct result of an unfortunate, high-profile event that attracts media attention. Agencies specifically focusing on predicting fatalities cited recent recommendations from the Commission to Eliminate Child Abuse and Neglect Fatalities (CECANF) as their primary reason for prioritizing elevated risk analysis. However, not every agency interviewed is focusing on predicting near-fatality or fatality events. While agreeing that a child fatality is the worst outcome for a child in their system, some felt that other, more prevalent problems may be better suited to a predictive analytics solution. Other elevated risk problems mentioned during interviews include identifying children at risk for re-abuse, likelihood of foster care placement, or future juvenile justice involvement.

Currently, many agencies utilize the National Council on Crime & Delinquency's Structured Decision Making® (SDM) model as a way to identify a child's risk for abuse or neglect in the future. The SDM model provides a methodology for a caseworker to gather and evaluate information to support their decision. This is a data-driven exercise to help create a defendable decision that optimizes resources. While the SDM model has the same goal of identifying children with an elevated risk of future abuse or neglect, it typically concentrates on the child's household or environmental characteristics to calculate risk. The predictive analytics solutions discussed in the context of this document can focus on a much broader list of characteristics, including the attributes of a particular child or multiple, unrelated perpetrators.

While much of this modeling is focused on the child as the unit of analysis, some stakeholders take different approaches to the problem. One jurisdiction is applying the models to the perpetrators, attempting to predict elevated risk that a perpetrator will abuse a child. However, in both situations, the models predict on the unique combinations of perpetrator and child. These options take the exact same outcome – a child fatality – and run a different analysis to answer a different question. Such differences will have implications for how the model results are used. In this specific example, answering the question 'Which children are at high risk to experience a fatality?' is operationally very different from the question 'Which perpetrators are at high likelihood to inflict fatal injury?' In both cases, whether focusing on the risk of a child being abused or the perpetrator offending, these predictive models are built upon a set of historical data and then applied to the unit of study in real-time to understand their current risk level. These models produce a "score" (typically a number on a scale of 0-100 or 0-1) that allows an agency to prioritize resources according to those children at highest risk. However, given the fidelity of the model or the underlying data feeding the model, these scores can be translated into a general categorization of high, medium, or low risk.

Repeated events

Predicting repeat events is the second most common problem that agencies are tackling with predictive analytics. There are two variants of this problem type. Under the first variant, the agency seeks to predict whether the child will experience (or the agency will receive) repeated maltreatment allegations over a certain time period. This problem can be focused on both children and perpetrators to children who are at a high likelihood of experiencing repeated maltreatment allegations during the interval, though models typically look at all of the data for a perpetrator or all of the data for a child in a singular capacity (i.e., one at a time) in order to tease out attributes of a particular person. In one agency, the agency's caseworker does not see a number, but rather a flag for that particular child indicating if the child's likelihood for repeat maltreatment exceeds a threshold predetermined by the agency.

The second type of child welfare event agencies are seeking to predict with this type of analysis is foster care re-entry and/or permanency. Predicting re-entry/permanency is often specifically focused on the child as the unit of analysis. For both of the repeated events problems, agencies typically cite a tragic incident that received media attention and a desire to impact the cyclical nature of child welfare and focus attention on the right cases as reasoning behind choosing to address these particular problems.

System issues

Problems that are categorized into system issues deal with the nature of the interactions between child welfare agencies and other service systems. These are often complex and interrelated situations, such as trying to predict which youth who age out of foster care have a high likelihood of becoming the perpetrator of child abuse/neglect later in life, or trying to estimate future child welfare caseloads within a region based on indicators related to parental opioid abuse. Such predictive models could enable child welfare agencies to improve outcomes by providing services to young parents who share characteristics of current perpetrators, but the modeling effort itself is quite complicated in bringing together multiple data sources and forecasting out years into the future. While these models are designed to be able to identify families in need of intervention, the corresponding intervention strategies often fall outside of the jurisdiction of the child welfare agency, to agencies/entities such as home visiting, early childhood, or adult behavioral health programs. Acting on these models requires additional collaboration and funding from another agency that may not share the same goals or have the bandwidth to accommodate child welfare clients and issues. Although these collaborations have the potential to enact long-term change by preventing potential maltreatment before it escalates, the implementation challenges can cause such system models to fall short in utility when compared to other predictive analytics solutions that can provide direct and actionable insights within the child welfare agency's influence.

One innovative approach uses a geographical, neighborhood-centric, lens – attempting to predict elevated risk of child abuse by people living in a ½-block area. This approach allows collaborative agencies to reduce risk in a particular region of their jurisdiction by identifying geographic areas to focus resources for maximum impact. While this approach is unique in identifying potential high risk neighborhoods or blocks, it can be difficult to determine the effectiveness of corresponding intervention strategies. In particular, there could be multiple changes to the surrounding environment that could lower the indicators for high risk of abuse areas, such as boosting the local economy or establishing community youth groups.

In addition to previously-mentioned reasons for settling on a particular problem, such as a desire to efficiently prevent tragic incidents, agencies implementing system issue predictive analytics mentioned a desire to implement proactive measures to combat child abuse (as opposed to reactive measures).

Agency operations

A final set of applications being explored within child welfare agencies relates to using predictive analytics as a tool for making resource management decisions. Predictive models addressing agency operations can allow an agency to utilize current caseload characteristics to forecast future caseworker workload or predict how much caseworker involvement a child will need to achieve an optimal outcome. This can help an agency make appropriate staffing decisions, such as when and how many caseworkers to hire by anticipating the incoming demand on their agency. In addition, an agency can effectively manage their caseworkers to create an environment that is less stressful, allowing them to manage an appropriate caseload and result in less caseworker attrition.

Agency operations issues often have different sets of motivators than the previous problems; agencies can be worried that other predictive analytics models – such as predicting child risk – will create significantly more work for the caseworkers. Modeling internal operations, including caseload size, may help to determine an optimal resource allocation that mitigates these potential concerns about implementing other types of predictive analytics solutions.

Ethical considerations and other constraints

Ethical controversy abounds in the context of predictive analytics applied to child welfare.[3] How can we preserve a child's privacy? How can we avoid unfairly profiling families? These are just some of the questions that arise when agencies discuss whether or not to use predictive analytics. Many agencies expressed trepidation about choosing certain problems to address via predictive analytics due to these ethical concerns. For example, child welfare agencies have no authority to intervene in a family's life if that particular family has not been reported to the child welfare system. Running large scale predictive models on every person and every family at birth – a commonly cited way to build a true proactive model that has the potential to prevent incidents before they escalate – is problematic if families do not voluntarily engage in preventive programming. A high risk score cannot directly lead to government intervention without a mandate for action. As such, many agencies are focusing their modeling efforts on the subset of the population actively involved in the child welfare system. However, this can result in creating a racial bias given the frequent overrepresentation of some minority families among those reported to CPS. For this reason, some agencies restrict the use of variables indicating a child or perpetrator's race in the modeling process.

The timeline for modeling these problems – whether re-abuse or elevated risk – is also subject to contention. One stakeholder prefers to model directly from birth, using information from birth certificates as the base set of information to inform the models about a child's parents, their health, and their home environment. The rationale for this approach is that this information is less subjective and provides visibility into the time in a child's life when he/she does not regularly interact with mandated reporters such as teachers or police officers. Other stakeholders contend that it is not ethical to predict on such information from a child who has yet to be abused, instead preferring to focus their modeling on children

[3] Katz, J., Sanders, D., Smith, S., & Specia, J. (2016, May 17). Leaders' Perspectives: Execution, Opportunities, and Challenges. Panel during American Enterprise Institute's 'Preventing Harm to Children through Predictive Analytics' event, Washington, D.C. [Transcript] Retrieved from https://www.aei.org/wp-content/uploads/2016/02/160517-AEI-Preventing-Harm-to-Children.pdf

who are already the subject of active cases. However, such interventions would be reactive instead of proactive and could reduce the overall impact of predictive analytics.

A third potential obstacle concerns access to data from agencies outside of child welfare's jurisdiction. In an ideal situation, agencies would also be able to predict the risk that a child will be involved in the juvenile justice system or develop substance abuse problems – or at the very least to include historical data on these problems into their modeling algorithms. However, the availability of cross-agency data is lacking and often impedes such analyses. For more discussion on data and its shortcomings, see Data section below.

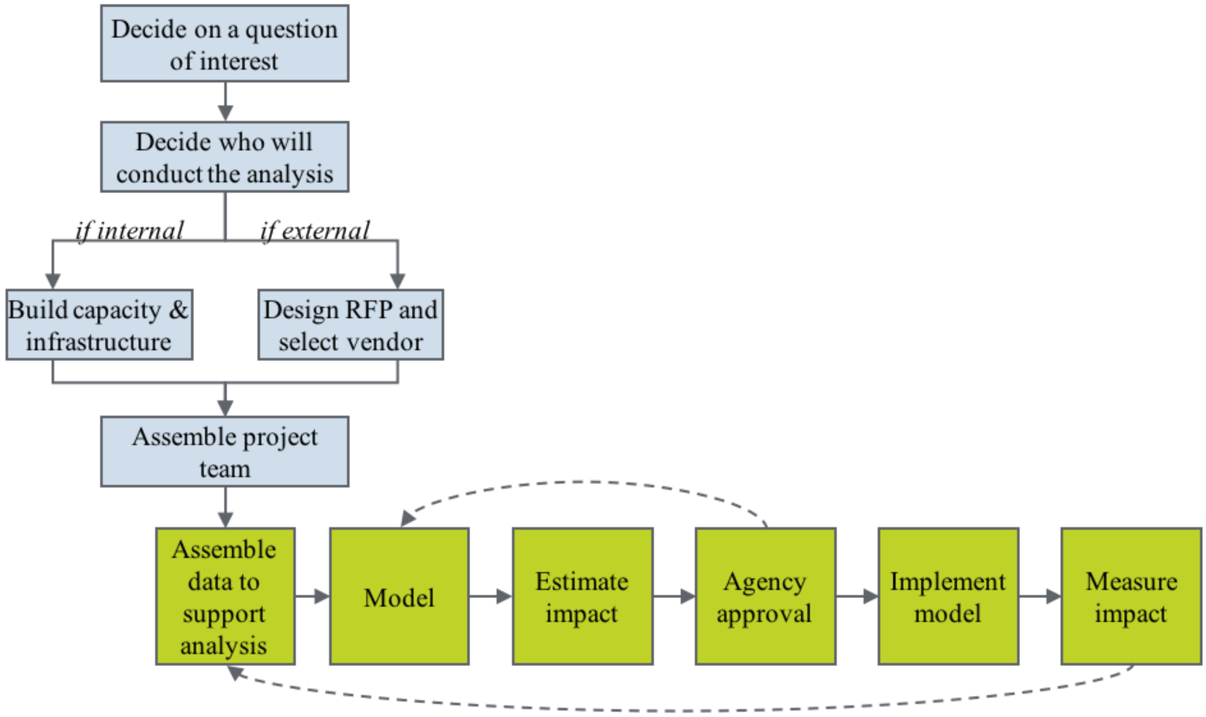
Every agency contacted was cognizant about the ethical concerns surrounding the use of predictive analytics in child welfare, such as violating privacy rights, misinterpretation of data, and building models that inadvertently incorporate demographic biases inherent to the data. Some agencies have addressed these concerns by providing transparency in the model building process or adding model results as another piece of information in a caseworker's decision making process, instead of replacing existing processes. Agencies have also responded to these concerns in their choices of the problems to address and how the predictive results are integrated into the conventional practice. Agencies choose their problems by determining which can be addressed most robustly given the available data and strategies for differential intervention. The stakeholders with whom we spoke largely concluded that concerns about data use in these applications should be taken seriously but should not wholly prevent the use of predictive analytics.

Consequently, many agencies seem committed to recommendations such those found in the CECANF report to integrate predictive analytics into child welfare processes. While four agencies did report that it took a bit of effort to explain predictive analytics and how the ethical concerns are being addressed to external stakeholders, such as the heads of other agencies and elected officials, they did not encounter significant pushback to the idea. To combat potential communication problems, some child welfare agencies started their implementation of predictive analytics with a small group of stakeholders devoted to the idea and expanded the group outwards, encouraging forward thinking beyond the crisis of the moment.

Predictive analytics solution design and construction

Once agencies determined the question or challenge they sought to address with predictive analytics, they started down a path to design a solution that could be integrated into their practice. While agencies have pursued different paths, these generally follow the approach outlined in Figure 1. This figure describes some of the key steps and decision points in an implementation of predictive analytics in child welfare. This process is broken into two parts: the first phase, shown in blue, represents initial activities necessary for the predictive analytics project and involves setting up resources. The second phase, in green, represents the process of model-building and implementation. The initial path through this process is sequential; however, the last part should be repeated over time to ensure model accuracy and effectiveness.

Figure 1. Flow chart of key decision points



Data

Data is the most crucial part of a predictive analytics implementation; without useable data, no further work can be conducted. Four of the agencies and all of the researchers interviewed agreed that acquiring data appropriate for use in a predictive analytics model was the most challenging aspect of the implementation, and that a range of barriers can prevent access to the necessary data. Of these key data barriers, the most prevalent were the ability to share data because of the statutes governing data sharing, and the interpretation of these statutes, as well as the quality of the data already available to each agency.

A number of federal and private activities create the context within which predictive analytics efforts are being explored in the child welfare field. At the national level, there has been increasing interest in the use of data and evidence for policy making and in ensuring federal data is available for improving programs. Most visibly, Congress in 2016 created the Commission on Evidence-Based Policymaking. The Commission is charged with developing a strategy for increasing the availability and use of data in order to build evidence about government programs, while protecting privacy and confidentiality. Through its work, Commission members are studying how data, research, and evaluation are currently used to build evidence, and how the government may strengthen evidence-building efforts.

Within the Administration for Children and Families, the component of the U.S. Department of Health and Human Services that operates most human services programs, the agency's Interoperability Initiative is making efforts to improve the ways human services agencies share data in order to better address clients' needs. Among their efforts has been to develop data exchange standards that improve the ability of human services agencies and their partners to exchange data with each other. Specific projects in the child welfare arena have enhanced information exchange to speed interstate adoptions, notify Indian tribes when Indian children are taken into custody by state child welfare agencies, and enable child welfare and child support agencies to share information with each other more efficiently.

Outside of government, Actionable Intelligence for Social Policy, based at the University of Pennsylvania, seeks "to improve the quality of education, health and human services' policies and practices through the use of integrated data systems."[4] In addition, Stewards of Change uses an interdisciplinary approach to bring the best business practices, models and tools to government and nonprofits through conferences, consulting, training and vision mapping. These tools focus on improving the use of data to drive practice and decisions and include specific efforts aimed at interoperability of data systems.

These varied efforts illustrate that challenges exist to integrating and using varied data sources for program purposes, including predictive analytics. Every agency brought up accessibility of data as a significant obstacle to implementing their predictive analytics approach. In an ideal world, predictive analytics projects would be based on the full spectrum of data available for a specific child or adult, including, but not limited to, data from child welfare (e.g., current investigation characteristics, previous investigation history, foster care/movement history, family dynamics), criminal records, substance abuse treatment histories, and health record systems. Currently, most predictive analytics models use the child welfare agencies' data in combination with other open source data – such as demographics, public crime data, and other community-level information. In many jurisdictions, criminal record data, substance abuse data, and health data are closed data sources owned by other government entities, and accessing that closed source information requires an agreement between the requisite parties. While not addressed specifically during the interviews conducted for this project, certain local-level agencies cannot legally share their data without changing governing state law, creating barriers to these efforts. In considering which barriers are most problematic, it is helpful to recognize that the list of necessary data sources is not the same for every predictive analytics model. Depending on the problem, some data sources prove to be more impactful than others. For example, a child's homelessness history might not be as useful to an elevated risk for abuse model compared to a model used to predict re-entry into foster care.

Even if it is legal for agencies to share data, political and privacy-related barriers often prevent such data sharing from occurring. The act of drawing up data sharing agreements between agencies is a nontrivial task due to differing opinions on what data can be shared, how it can be used, and how it should be stored. Of the agencies we spoke with that have started down this path, the average time to the final assembly of a data warehouse[5] is more than two years. Final assembly of a data warehouse includes pulling together the necessary hardware, raw data feeds after agreements have been signed, and creating the ability to

[4] About Us – Actionable Intelligence for Social Policy (2017). https://www.aisp.upenn.edu/about-us/

[5] A data warehouse is an electronic storage architecture designed to hold data extracted from transaction systems, operational data stores, and external sources. The warehouse then combines that data in an aggregate, summary form suitable for enterprise-wide data analysis and reporting for predefined business needs. Gartner glossary

join data together for analysis. In their article about data sharing across states, Green et.al. concluded the following, which echoes the sentiments that agencies relayed about data sharing *within* their own state:

> "One of the key lessons for researchers who seek to access child welfare data across multiple states is to plan for the amount of time needed to develop and finalize agreements, and to receive data from state agencies."[6]

Jurisdictions often have different governing statutes – and different interpretations of those laws – concerning both the definitions of personally identifiable information (PII) and protected health information (PHI) and requirements about how to preserve the safety of that information. These agencies will typically then try to compromise by providing de-identifiable data for analysis. However,, while potentially workable for some purposes, de-identified data cannot address the need for granular data that enables matching across systems and encourages successful analytic projects. As a result of the issues surrounding data sharing agreements, three of the child welfare agencies whose staff were interviewed are currently running predictive analytics using only their own data, with future plans to incorporate more detailed data as agreements are finalized. Furthermore, to be able to effectively implement predictive analytics, these data agreements need to result in real-time data sharing environments such that model scores can be generated based on the most up-to-date data. Two child welfare agencies commented on how historical snapshots provided once every few months would not be useful for programmatic interventions but could be useful for research purposes where the timing of obtaining results is not as important as the specific result. For instance, if key factors in your model include a new arrest for someone in the household or a stay in a homeless shelter, data with a 3-month lag is insufficient when applied in a real-time solution. These agencies also mentioned that real-time data, or at least a relatively frequent update period, is needed to implement models to be used in case practice, which is the goal for every agency interviewed.

Regardless of the status of data sharing agreements, all agencies mentioned data quality as a significant issue when beginning a predictive analytics project. Child welfare agencies cited significant concerns about the reliability of their own data, repeatedly mentioning the need for ongoing entity resolution exercises to remove duplicate people from their databases, in addition to implementing data quality checks as information is entered into the electronic systems. Entity resolution can be a significant advanced analytics project in itself, particularly if the data lack unique identifiers, such as a social security number, to enable the efficient and accurate consolidation of records. In the absence of a unique identifier across data sets, advanced modeling techniques often need to be employed to compute a statistical likelihood that multiple records are, in fact, relating to the same person.[7] These techniques, often referred to as entity resolution or record linkage, are not within the scope of predictive analytics but require a non-

[6] Green, B. L., Ayoub, C., Bartlett, J. D., Furrer, C., Von Ende, A., Chazan-Cohen, R., … Nygren, P. (2015). It's not as simple as it sounds: Problems and solutions in accessing and using administrative child welfare data for evaluating the impact of early childhood interventions. *Children and Youth Services Review*, *57*, 40–49. doi:10.1016/j.childyouth.2015.07.015

[7] Discussion on entity resolution adapted from Gartner, Inc. (2016). Gartner glossary

trivial amount of resources and if not conducted at an earlier stage, are required before utilizing multiple data sources for a predictive analytics project.

While data sharing agreements and data quality are two separate issues with separate solutions, in many cases, child welfare agencies who are leading in establishing data sharing agreements are also further along in rectifying quality issues. Those child welfare agencies recognized that outside data sources could be valuable for their predictive analytics projects, but recognized that their own internal data quality was not good enough initially to combine with those external data sources as they started predictive analytics projects.

Analysis sourcing

Child welfare agencies currently pursuing predictive analytics projects expressed a range of opinions about whether, ultimately, they wished to conduct their analyses in-house or through vendor contracts. However, agencies universally recognized that they do not currently have the staffing capacity to make in-house analysis feasible. Predictive analytics requires comprehensive training in data science and statistics, and, at the time of writing, the field is in high demand and low supply across a wide variety of industries. Child welfare agencies often report that they not only lack the appropriate skills in their current staff, but also the necessary budget to hire new staff with the requisite backgrounds at a competitive salary for the industry at large.

Given that agencies agree that they cannot implement predictive analytics with their current staff, every agency with which CAMH spoke is using some form of outsourcing and contracting to build at least the initial set of predictive models and get their projects off the ground. However, the agencies split evenly on what type of vendor they preferred: a nonprofit/academic researcher partnership or a for-profit company. The agencies who favored partnering with nonprofits or academia cited the unbiased nature of the vendor as a large selling point; they viewed these vendors as 'in the game for the right reasons' and would attempt to achieve the best results without the objective of making profit off the engagement.

Other agencies, however, chose for-profit companies to build their predictive models, citing their breadth of experience implementing such predictive analytic projects across a wide variety of industries, as well as within child welfare itself, as a large draw. For-profit companies also have the advantage of being able to better integrate with business/agency operations and provide a highly-polished, implementation-ready final product. While these agencies tackle similar problems, each predictive analytics approach still requires a significant amount of customization due to the lack of standardization of data and the differences in each jurisdiction's processes. Until this issue is resolved, for-profit companies cannot market an effective turnkey solution. Of the interviewed agencies, those that partnered with for-profit companies have less visibility into the work that is being done and less of a feeling of ownership over the finished product. None of the agencies contacted could describe details on cost and performance because the agencies are involved in an ongoing contract with the vendor or have recently implemented the software products and do not yet have a sufficient track record from which to assess costs.

Despite the challenges of developing predictive analytics capabilities in the government, two of the agencies interviewed are attempting to build capacity in-house to be able to create, update, and implement predictive analytic models once an outside vendor has developed the initial model. In these situations, the agency partners with an external vendor to build the initial set of models and get the program off the ground. Following the initial launch, they can train and educate their staff alongside the contracted partners, while building the next round of models. This approach allows agencies to build up skills by both learning from the external partners and having more time to hire qualified in-house staff. In such cases, the agency noted that their staff members do not have the requisite skillsets at the beginning of a project, and it would be better to work alongside experts to implement predictive analytics sooner, rather than building internal capacity for a decade and then starting the project. Some jurisdictions are taking this capacity building one step further by building partnerships with their local universities that will create programs to give students the necessary skills upon graduation. The Florida State Institute for Child Welfare[8] is one such partnership intended to foster data science skills through academic programs.

Other technical obstacles

While accessing data and assembling a team to do the analysis were the two most commonly cited pressure points for implementing analytic projects, agencies also identified other obstacles. When asked about their timeline to implement predictive analytics projects, every agency reported a multi-year planning process that preceded the modeling efforts begun in the last 18-24 months. Some agencies issue shorter-term requests for proposals (RFPs) that use existing data warehouses and encompass options to extend, such as Allegheny County's 2014 Decision Support Tools and Predictive Analysis RFP.[9] Other agencies enter into longer-term partnerships to help clean the data and build predictive analytics capacity. In both cases, however, the total time spent on the implementation of predictive analytics from the first effort is multiple years. Expecting to go from RFP to implemented model in a year or less is likely unrealistic as expressed by the interviewed child welfare agencies and their partners.

Approximately half of the agencies interviewed for this report mentioned cost barriers as a key challenge. Furthermore, some of the agencies that mentioned cost were the same agencies that had not yet invested in a centralized, cleaned data warehouse. The incremental cost of a predictive analytics project is much smaller when the child welfare agency already has invested funding to clean and organize data into a data warehouse that can support analytics projects. While predictive analytics may not be an exorbitant cost to agencies, there is limited surplus funding in agency budgets to spearhead a task that is not seen as essential to the day-to-day operations of the agency. As a result, two agencies interviewed have sought outside grants to fund their predictive analytics work.

[8] Florida Institute for Child Welfare. http://csw.fsu.edu/ficw/

[9] Allegheny County Department of Human Services. (2014, February). *Decision Support Tools and Predictive Analytics in Human Services RFP*. Retrieved from http://www.alleghenycounty.us/Human-Services/Resources/Doing-Business/Solicitations/Archive.aspx

Implementation, impact, and response

Each child welfare agency had the goal of implementing predictive analytics solutions into their day-to-day practice or process. Similar to any other predictive analytics implementation, each change to the process should be implemented in a way that allows for a child welfare agency to measure the impact. Ideally, each predictive analytics project would result in a measurable influence on the child welfare system and provide justification for the investment in predictive analytics.

Integrating models into existing processes and measuring results

Three of the child welfare jurisdictions interviewed are still in the early phases of model development: gathering data and running preliminary models. It was often reported that the most challenging part about implementing predictive analytics in child welfare is building the analytic models into the day-to-day practice, in such a way that does not burden caseworkers but encourages meaningful change. In a discussion paper presented at the 2015 NDIS New World Disability Conference in Brisbane, Brian Lee-Archer, Thomas Boulton, and Kylie Watson discuss this issue:

> "The challenge going forward is akin to finding an appropriate way to integrate the analogue world of professional judgment, experience and interpretation with the digital world of predictive analytics."[10]

Of the agencies interviewed that have taken that next step toward implementation, two are integrating the predictive analytics scores as added pieces of information in the child welfare toolkit. Risk scores are often displayed as a thermometer which obscures some of the detailed prediction, attempting to highlight that models are never perfect. Some agencies plan on allowing caseworkers access to the detailed scores, and others plan on restricting the use of risk scores for screening purposes in the call center that receives child abuse and neglect reports. Furthermore, in an attempt to prevent potential bias, some agencies are requiring that the screener calculate and enter his/her own risk scores based on professional judgment or other tools before gaining access to the model risk score. Given that many jurisdictions are just now starting to think about the implementations of predictive analytics, it is hard to say what unifying practice on implementation will coalesce, if any, let alone whether any will be recognized as the best practice.

Of the child welfare agencies that have developed models and integrated the model results and predictions into their day-to-day operating processes, none were willing to comment on the success or impact of the initiatives, citing that not enough time has passed since implementation to have confidence in such measures. Nevertheless, the agencies report that preliminary implementations have not raised any red flags for future success or impact. Academics and other independent experts have recommended that looking at certain metrics – such as over/under intervening, the rates of false positives/negatives, and the total number of allegations accurately resulting in an open case – could be useful methods of assessing the effectiveness of predictive analytics. Many agencies do have plans in place to implement

[10] Lee-Archer, B., Boulton, T., & Watson, K. (2015). SOCIAL INVESTMENT IN THE DIGITAL ERA: Improving social outcomes through predictive analytics and managing the emerging moral hazard and ethical challenges. In *SAP Institute for Digital Government*. Brisbane: NDIS New World Disability Conference.

reviews of the models and risk factors as they collect more data about the use of the models. They will then conduct regular, recurring reviews and revise the predictive analytics as needed – both the models and the manner in which they are implemented. In the absence of a scientifically controlled experiment, many agencies are wary of claiming that any one result is a direct consequence of predictive analytics. Child welfare agencies are constantly making changes to their policies and procedures, and predictive analytics are just one of those changes.

Post-implementation concerns

Some common themes emerged from the stakeholder interviews about concerns that may arise post-implementation, including ethics, caseworker workload, and staffing. While ethical concerns often rise during the problem assessment phase of predictive analytics, they come up again during implementation and response: how can risk scores best be displayed to caseworkers to convey uncertainty in the assessment? How can the nuance of the modeling algorithm be conveyed? A common theme was the need for significant training for whomever will be interacting with the risk scores. With these new tools, caseworkers and screeners need to understand how the score is constructed and what it means before being able to effectively use it to influence their casework practice. In this regard, easy-to-explain modeling algorithms and displaying risk scores as larger categories (e.g., high/medium/low risk) attempt to bridge the gap.

Similarly, agencies struggle with how to best combat the trend to see somewhat high risk scores and automatically create a case for that child, without regard for caseworker judgment or workload. Alternatively, agencies could encounter a situation where universally low risk scores could result in not enough allegations being elevated to case status. Many agencies work with unionized caseworkers and operate under contracts that specify specific rules and workload requirements as well as penalties for adjusting the contract. While caseworkers generally agree that predictive analytics can be a useful tool in preventing abuse or neglect, changes to caseworker workload as a result of the implementation of predictive analytics can open the door to union review. This is a significant concern in jurisdictions that have already implemented predictive analytics, and one child welfare agency noted that caseloads tend to increase after the implementation of predictive analytics unless mitigation strategies are in place (e.g., training on the uses of the risk score or using the model to prioritize resources instead of automatically assigning cases).

Lastly, implementing predictive analytics can lead to additional concerns around the sustainability of such efforts, specifically around the ability of child welfare agencies to continually monitor and maintain the models. Statisticians agree that a statistical model should be updated with regular frequency, although the time between updates can vary depending on the model.[11] A model update can fall into two categories: recalibrating the features used in the model and rebuilding the model. In most cases, recalibrating the model requires less effort and therefore can be done more frequently than fully

[11] Sculley, D., Holt, G., Golovin, D., Phillips, T., Davydov, E., Ebner, D., Young, M. (2014). Machine learning: The high interest credit card of technical debt. *Google Research*. Retrieved from [google research]

rebuilding the predictive model. Agencies generally recognize that a model needs to be continually refreshed (at least annually) with updated data; a predictive analytics project is a cyclical project that is never truly done. As a result, agencies potentially face longstanding contracts with external vendors or the need to build internal capacity. Without knowing the actual return on investment, every child welfare agency expressed concern at the initial base cost for model development and the necessary long-term investment, from both a monetary perspective and a time perspective.

Other benefits from implementing predictive analytics

One of the largest, most significant benefits of a predictive analytics project in child welfare is that such a project forces an agency to examine and improve its data. A common theme is that child welfare data is messy and inconsistent, particularly the fields that are not part of nationally consistent federal reporting through the Adoption and Foster Care Analysis Reporting System (AFCARS) or the National Child Abuse and Neglect Data System (NCANDS). In implementing predictive analytics, child welfare agencies are forced to clean up their databases, resolve duplicate entries, identify inconsistencies, and implement data quality assurance processes such that future data is much more reliable. This investment in data quality not only helps future predictive analytics projects, but also enables more advanced projects in all types of analytics. Additionally, work done to establish data warehouses and/or data sharing agreements across different agencies also helps improve data quality. Some agencies may take advantage of new regulations regarding development of Comprehensive Child Welfare Information Systems (CCWIS) to upgrade their case management information systems to support their child welfare program needs, including their ability to exchange data with other state programs. These improved data systems may be established in ways that improve states' capacities to implement predictive analytics projects.

Next steps: what does the future hold?

Predictive analytics as a field is still evolving in its application to child welfare. Key to the implementation of a predictive analytics solution is the ability to provide sustained, long-term improvement to the appropriate outcome. The goal of any predictive analytics effort is, and should be, to improve the trajectories of the families in the system.

In a discussion paper applying a specific predictive analytics algorithm to child welfare, Kum et.al. conclude that the key factors to a successful predictive analytics implementation include:

> "trust, real support through policies and funding, access to good technical expertise in both the content area and information technology, and training."[12]

Child welfare agencies are looking to continue their efforts in improving outcomes via support from predictive analytics. While some solutions are currently in place, there are many solutions still in their

[12] Kum, H.-C., Joy Stewart, C., Rose, R. A., & Duncan, D. F. (2015). Using big data for evidence based governance in child welfare. *Children and Youth Services Review*, *58*, 127–136. doi:10.1016/j.childyouth.2015.09.014

infancy. Solutions in place will need to be updated and maintained to ensure the validity of the model going forward in addition to staying current with any technology advancements.

Predictive analytics success stories

The limitations of predictive analytics withstanding, some jurisdictions have had success in implementing applications of predictive analytics. In the long term, a successful predictive model will have a measurable impact on the lives of children and families in the agency's jurisdiction by using the model results to enhance policy and practice. Ongoing outcomes monitoring can help to evaluate these efforts and assist states in achieving positive results. However, given that child welfare agencies are early in their implementation of predictive analytics, the examples presented here define success as the development of a predictive analytics model intended for production that is based on one or more datasets – essentially taking successful steps towards that end goal.

One agency we spoke with is in the very early stages of implementing a set of predictive models that crosscuts the problem categories mentioned in this document. The models are in their infancy, but strong partnerships and institutional support – in addition to a developed and maintained data warehouse – help to lay the groundwork for this process. Another agency devoted a significant multi-year effort to building a comprehensive data warehouse that encompasses a wide range of health and human services data and has used that information to build predictive models. Again the models are in their infancy, but the agency has a strong data foundation on which to build future efforts. Across the board, even if agencies do not have a production-ready model in development, all agencies have successfully cleaned up their internal datasets and started to implement policies to better ensure that future data is more robust. On the researcher side, some analysis using neighborhood-based risk modeling has led to the identification of community-based prevention and family strengthening programs, and other researchers have started to integrate data across agencies. While none of these efforts are a fully-fledged implementation of predictive analytics, they are all taking successful steps towards the ultimate goal of improving a child's life through predictive analytics-based interventions.

Federal opportunities to guide the field

Throughout the discussions with various organizations and researchers, CAMH learned of ways in which state and local-level agencies suggest the federal government could support them as they look to implement predictive analytics. These include the following:

- **Provide Sample Data Sharing Agreements and RFPs.** One opportunity would involve providing child welfare agencies with template legal agreements in order to decrease the time for these agreements to be established. These legal agreements could include data sharing agreements between local-level agencies. Agencies would also like documentation and technical assistance as they develop RFPs for hiring outside vendors to develop predictive analytics solutions, helping them to be informed consumers and increase the likelihood of a successful acquisition.

- **Offer Technical Assistance on Data Issues.** While AFCARS, NCANDS and data from states' SACWIS/CCWIS provide a good baseline for data collection and reporting, it is often not a complete source of data needed for a predictive analytics solution. Respondents said that they would like guidance on the additional data elements that could be added to those currently being captured in states' data systems. This would provide agencies with an understanding of what data they would need before considering predictive analytic projects. These data sources could be defined by a common schema, allowing for jurisdictions to create consistent definitions for specific fields and their values as recommended in the CECANF final report.[13]

- **Develop a Programming Library**. Respondents also thought that support for local-level agencies would be helpful in the development and dissemination of sample predictive analytic scripts or programs for the common problems mentioned in Table 1 above (i.e., elevated risk during or following preventive services, repeated events, system issues, and/or agency operations). These scripts could provide a baseline from which agencies can start and see what potential impact a predictive analytics solution might have within their agency. These scripts could also define specific metrics for assessing a predictive model and highlight any corresponding concerns an agency might have before implementation.

- **Create a Central Data Repository and Analytic Environment.** Another option is to allow child welfare agencies across the country to upload data to a secure central repository and analytic environment hosted by a non-governmental trusted third party. This would enable an agency to understand how its performance and outcomes compare to the average performance of other similar child welfare agencies across the country, providing a relatively safe space in which to use data in new ways with less fear that findings might be made public or used in litigation. The trusted third party could provide the analytical expertise for analyzing the data provided and disseminate the corresponding analyses through federal government communication channels. This would also allow the trusted third party and agencies to analyze sensitive questions about potential deficiencies or weaknesses and identify areas for improvement in a secure and improvement-focused environment. For example, an agency could see which characteristics are most important in understanding why a child fatality might occur within its jurisdiction. In cases of analyzing metrics of rare or unlikely events within a single jurisdiction, such as a child fatality, this repository would be able to bring together enough of those rare data points in order to identify any potential trends or commonalities between events. A collection of those rare events would allow an agency to form a basis for a predictive model tailored towards the unique characteristics of its jurisdiction. Such a system might be able to provide a networking capability for agencies to see where predictive analytics are successfully addressing certain problems. This option is based, in part, on the success of the model employed by the aviation safety community to integrate sensitive data and develop analytical tools that have led to remarkable improvements in aviation safety.[14] Another successful model is the KIDS COUNT data center, a project of the Annie E. Casey Foundation that hosts hundreds of indicators and allows users to create reports and graphics to support smart decisions about children and families.[15]

---

[13] Commission to Eliminate Child Abuse and Neglect Fatalities. (2016). Within our reach: A national strategy to eliminate child abuse and neglect fatalities. Washington, DC: Government Printing Office.

[14] Rosenkrans, W. (2011, November). No turning back. AeroSafetyWorld. Retrieved from https://flightsafety.org/asw-article/no-turning-back/

[15] KIDS Count Data Center from the Annie E. Casey Foundation (2017). Retrieved from: http://datacenter.kidscount.org/

- **Find Ways to Support the Development of Local Technical Expertise**. As mentioned above, building a team of staff with the ability to perform and maintain predictive analytic solutions has proven difficult for two of the agencies interviewed that tried to develop this capacity in house. These agencies would welcome federal support, financial and otherwise, in order to build up this capability within their agencies. It could be useful to seed the field and provide base resources for state and local-level agencies to build an initial predictive analytics capacity. The federal government could also provide or facilitate a training program in predictive analytics to help emphasize a baseline skillset for staff within child welfare agencies.

Conclusion

Predictive analytics is a new tool being used to support child welfare practice. Child welfare agencies and researchers have started looking at many ways to improve upon their current practices through the application of predictive analytic solutions into their process. In many cases, these approaches provide another data point into the decision making process for caseworkers. In addition to supporting caseworkers, agencies are using predictive analytics as a way to reform their thinking around intervention strategies, often across government agencies and stakeholder groups that would allow for more collaboration to achieve the goal of providing a safe and supportive environment for children.

However, for the agencies interviewed, the path to using predictive analytics has not been straightforward or easy. Agencies remain divided on the optimal way to approach researching predictive analytic solutions and, indeed, different solutions may be optimal depending on agencies' circumstances. Child welfare agencies express differing opinions as to whether it is best to hire external vendors or to partner with an academic institution or a not-for-profit organization. There is a common challenge in both cases though: data. Each agency commented on the challenges including setting up data sharing agreements, the ability to merge different data sources, and the overall quality of data in being used for predictive analytics.

Predictive analytics is still in its development or initial implementation phase at many organizations. The knowledge learned through the process of developing predictive analytic solutions can be helpful for other agencies considering a similar approach to their problems. However, the federal government can assist in a few ways to decrease or eliminate some of the common challenges agencies can face by providing model documents, programming scripts, guidance and other support to make predictive analytics more accessible to a wider range of agencies and help them build their internal capacities to use data in this manner.

Interviews indicated that some agencies may have unrealistic expectations about the near term value and impact of predictive analytics. Predictive analytics alone will not provide a panacea for all challenges child welfare agencies face but can be an important tool to support processes, practices, policies, and other systemic changes and improvements. This is a phenomenon that is common when new technologies or solutions are thought to hold great promise[16]. There are certainly a number of success stories from child welfare agencies. However, due to the challenges outlined above, there is likely to be an extended learning process before agencies mature in their understanding of the strengths and limitations of

[16] Gartner Hype Cycle (2016). gartner.com/technology/research/methodologies

predictive models and how to best incorporate them into the other tools, strategies and interventions needed to improve child welfare outcomes. Once those challenges have been worked through, predictive analytics can be a useful tool to help child welfare agencies to be better informed about the risks facing the families and children they support and to be better allocate limited resources to achieve the greatest impact.

Appendix

Detailed list of problems being modeled

Entity Identifier	Problem being modeled	Primary Data Source	Project Team	Current Status
A	Risk of fatality/near fatality	• Child welfare data sources[17] • Substance Abuse • Education • SNAP • Medicaid	Agency partnering with nonprofit external partner	Research Pilot
B	Risk of fatality/near fatality	• Child welfare data sources[17]	Agency partnering with for profit external partner	Research Pilot
B	Repeat reporting/maltreatment	• Child welfare data sources[17]	Agency partnering with for profit external partner	Research Pilot
C	Risk of fatality/near fatality	• Child welfare data sources[17] • Public health • Mental health	Agency partnering with for profit external partner	Research Pilot
D	Risk of a child welfare event at the time of a hotline call	• Child welfare data sources[17] • Substance abuse • Homelessness records • Criminal justice records • Education • Public housing • Location characteristics	Agency partnering with nonprofit external partner, academia	Production
E	Repeat reporting	• Child welfare data sources[17]	Agency partnering with nonprofit	Production

[17] Child welfare data encompasses data owned by the child welfare agency pertaining to the children, their families, and the corresponding events logged in the system (e.g., referrals, cases, foster care)

Entity Identifier	Problem being modeled	Primary Data Source	Project Team	Current Status
		• Location characteristics	external partner, academia	
E	Elevated risk during or following preventive services	• Child welfare data sources[17] • Location characteristics	Agency partnering with nonprofit external partner, academia	Production
E	Re-entry	• Child welfare data sources[17]	Agency partnering with nonprofit external partner, academia	Production
E	Intergenerational Involvement	• Child welfare data sources[17]	Agency partnering with nonprofit external partner, academia	Production
F	Community-based high risk for maltreatment fatality	• Child welfare data sources[17] • Community-based crime data • Location characteristics	Academia	Research Pilot
G	Risk for future child welfare involvement	• Hospital/birth records	Academia	Research Pilot

Acronyms

Acronym	Definition
ACF	Administration for Children and Families
AFCARS	Adoption and Foster Care Analysis and Reporting System
ASPE	Assistant Secretary for Planning and Evaluation
CAMH	CMS Alliance to Modernize Healthcare
CCWIS	Comprehensive Child Welfare Information System
CFSR	Child & Family Services Review
CECANF	Commission to Eliminate Child Abuse and Neglect Fatalities
CMS	Centers for Medicare & Medicaid Services
FFRDC	Federally Funded Research and Development Center
FOIA	Freedom of Information Act
HHS	Health and Human Services (e.g., Department of)
NCANDS	National Child Abuse and Neglect Data System
PHI	Protected Health Information
PII	Personally Identifiable Information
RFP	Request for Proposal
SACWIS	Statewide Automated Child Welfare System
SDM	National Council on Crime & Delinquency's Structured Decision Making®

Discussion schedule

Below is the list of agencies, researchers, and vendors that provided information to CAMH and was incorporated into this document. In addition to these entities that provided information for this document, a number of other potential informants – primarily private vendors – declined to participate.

Date	Entity Name	Entity Type
9/9/2016	San Diego Health & Human Services Agency	Child Welfare Agency
9/13/2016	New York City Administration for Children's Services	Child Welfare Agency
9/14/2016	Florida Department of Children and Families	Child Welfare Agency
9/14/2016	Andy Barclay	Nonprofit/Independent Expert
9/16/2016	Los Angeles County Department of Children and Family Services	Child Welfare Agency
9/19/2016	Allegheny County Department of Human Services	Child Welfare Agency
9/20/2016	Casey Family Programs	Nonprofit/Independent Expert
9/26/2016	Dr. Dyann Daley, Dr. Michael Bachmann	Academic Researcher
9/27/2016	Dr. Emily Putnam-Hornstein, Dr. Regan Foust	Academic Researcher
9/29/2016	Dr. Rebecca Orsi	Academic Researcher
9/30/2016	Will Jones (SAS)	For-Profit Company
10/27/2016	Dr. Jesse Russell	Nonprofit/Independent Expert