

Principles for Predictive Analytics in Child Welfare



Authors and Acknowledgments

Authors

Chris Scharenbroch, NCCD

Kathy Park, NCCD

Kristen Johnson, Ph.D., University of Minnesota, Department of Family Social Science

Acknowledgments

NCCD wishes to thank Tashira Halyard, Senior Associate at the Center for the Study of Social Policy and head of the Center's Alliance for Racial Equity in Child Welfare, for her collaboration and her contributions to this paper.

About NCCD

NCCD promotes just and equitable social services for individuals, families, and communities through research, public policy, and practice. Visit www.nccdglobal.org to learn more.

Introduction

The advent of “big data” and predictive analytics in the context of human services, and more specifically in the child welfare field, has prompted a burgeoning interest among child welfare system administrators, policy makers, and technology vendors in making such approaches the panacea for any potential abuse- and neglect-related outcome for children and families. The desire for solutions has spurred a handful of product development efforts by various private corporations and within academia aimed at demonstrating proof of concept, as well as some first attempts at implementation.

At the NCCD Children’s Research Center, we know that harnessing the power of data has the capacity to drive better outcomes not only for children and families served by child protective service systems, but also for the people on the front lines within these systems and for policy makers who impact how these systems are defined, resourced, incentivized, and measured. At the same time, we believe it is imperative that development and innovation using predictive analytics take place



We believe that all predictive models must offer a demonstrable improvement upon current system responses for children, families, and staff, given the tradeoffs in transparency and access to personal data these solutions require, and the resulting implications on equity and individual civil liberties.

in the context of explicitly defined values and principles.

General consensus on a “do no harm” framework around the use of predictive analytic models—a necessary baseline—currently exists in the child welfare field. However, we want to push the field to think and act beyond this standard. We believe that all predictive models must improve upon the current system response for children, families, and staff. This value is ultimately connected to the evaluation of success at promoting positive outcomes by system interventions, services, and programming. Researchers, practitioners, and leadership must consider the following principles during predictive model development, testing, implementation, ongoing evaluation, and continuous quality improvement.

Core Principles for Building and Applying Models

First, define the question; then, determine if a predictive analytics model is the right tool to help answer it.

The research or policy question that outlines the outcomes of interest and potential application of the predictive analytics tool is critical for appropriately identifying the sample and analysis methods. The question also informs evaluation criteria for tool testing and application. The problems of approaching the development of

Mathematical models always have exceptions and cannot account for the rare, clinical risk factors that a professional investigating worker can observe.

predictive models in the absence of a clearly defined question (i.e., what is the problem we are trying to solve?) are many, with the most concerning being the misapplication of model outputs in unintended and unanticipated ways that can potentially drive more people into system contact and exacerbate disparities (racial, ethnic, gender, socio-economic status, sexual orientation, etc.). For example, in child welfare, identifying someone as being at high risk of committing a child fatality without having any legal basis for government intervention has significant

implications for individual civil liberties and necessitates a critical discourse on the role and scope of statutory child protection. An important safeguard to ensure ethical and appropriate use of predictive analytics is to first fully define the question that needs to be answered, and then explore whether predictive analytics might be a useful tool in answering it. It is also useful to explore whether existing tools and methods with demonstrated effectiveness could be applied to the question/problem at hand.

Use model outputs to drive supportive (not punitive or net-widening) interventions.

Research documents the limitations and potential of child welfare administrative data (English, Brandford, & Coghland, 2000; Garnier & Poertner, 2000; Green et al., 2015). To date, predictive analytics tools developed from administrative data are useful, though not often very sensitive or specific. In addition, mathematical models always have exceptions and cannot account for the rare, clinical risk factors that a professional investigating worker can observe. Given the limitations of predictive modeling methods, the ethical issues surrounding identification of families with a high risk classification, and the potentially negative consequences of a false positive identification, model applications should be limited to preventive approaches and implemented if risk of unforeseen negative consequences is minimal. Examples of this include focusing provision of ongoing supportive child welfare services or targeting limited respite care resources to foster care providers who have higher likelihoods of placement disruptions.



Acknowledge that racial and ethnic bias is an issue.

The over-representation of children and families of color in the child welfare system at multiple decision points from intake through permanency results from institutional, systemic, and individual biases. Regular review of what type of data is going into predictive analytics approaches, how and when those data are being entered, and consistency across users who are entering data may, over time, help to mitigate some of the inequities represented in case management data. In the meantime, racial and ethnic disparities in the underlying data should be acknowledged rather than denied; predictive analytics tools should be used to classify individuals by the likelihood of experiencing subsequent system involvement for the purposes of supportive interventions rather than punitive system responses.

Further, predictive models help us understand how disparities in system involvement are often a function of ecology. Individuals and families who have higher likelihoods of experiencing system involvement are often clustered by certain characteristics such as location. These models can be a useful resource for targeting community-based prevention and service efforts.

Predictive models help us understand how disparities in system involvement are often a function of ecology.

Be transparent about the research, model, implications, and limitations (e.g., limitations of data sources, methods of analyses).

Workers will more likely find results valid and useful if they understand the predictive analytics model and how classifications and recommendations are derived. Modeling, statistics, and big data are new concepts to some people, and discomfort with mathematical algorithms can lead to anxieties about data accuracy and use. It is important to be transparent with agency workers, supervisors, and administrators about how data are used because data and model accuracy and reliability depend on the quality of staff's assessment findings and their recording and documentation of case actions.

Families served by the child welfare system have the right to information and understanding about how decisions are being made about them and their children.

The principle of transparency extends beyond internal agency decision makers and actors and should include key community and system stakeholders as means to garner critical buy-in and trust. Most importantly, the families served and impacted by the child welfare system should have the right to information and understanding about how decisions are being made about them and their children, given the high-stakes nature of child welfare decisions. Algorithms can be very successful in predicting likelihoods when

constructed from accurate data, and classification results are more likely to improve the effectiveness of practice when staff understand how and why the algorithm reached its result.

Support appropriate use and implementation of the tool with initial and ongoing testing and evaluation.

A critical part of ensuring that use of predictive analytics has the intended impact on practice—and no unintended impact—is monitoring and evaluating how the resulting information is used. A comprehensive evaluation will include (a) monitoring fidelity of model applications and their associated interventions to ensure that they match their design and intended use; (b) tracking changes in outcome to understand the impact and effectiveness of interventions and changes in practice that result from model implementation; and (c) conducting a study to measure the relative costs associated with model applications. Examining fidelity is key to determining whether model results are useful to staff and whether staff are following guidelines for implementation.



Guidelines for Predictive Analytics Development and Testing

Make sure the sample is appropriate given your question and proposed model use.

The sample should consist of a cohort of individuals or families that are representative of the population for which staff will be applying the predictive model. The sample must also support understanding of subgroups of the population (this would include sufficient data to examine

Predictive models should demonstrate validity and equity for the local population.

race/ethnicity, gender, age, location, etc.). Models should aim to leverage available data, including demographics, client characteristics, service history, history of system involvement, assessment information (e.g., risk, needs, case plans, etc.), and/or cross-systems data to describe samples with as much detail and specificity as possible. It is important to:

- Know the limitations of a sample (data quality, accuracy of data entry, when staff/workers complete data elements relative to field data collection and decision making).
- Understand the effectiveness of services. The ability to measure the effectiveness of services and programming is helpful to understanding how a model may operate in practice. For example, a model may use predictive factors

to identify groups of individuals with similar likelihoods of experiencing subsequent system involvement. In many cases, some of these individuals will have received services or programming while others will not. Comparing how similarly situated individuals experience subsequent system involvement by the services or programming may be useful in understanding program eligibility criteria and other practice considerations.

- Validate predictive tools on a population different than the one used to construct it. Classification results will be the most robust for the sample from which the assessment was constructed, and validating a risk assessment on a separate population provides a better approximation of how a risk assessment will perform when implemented (Altman & Royston, 2000).

Examine model performance for validity and equity.

Predictive models should demonstrate validity and equity for the local population either prior to implementation or as part of a pilot test. Ideally, multiple measures are used to compensate for the limitations of any one measure.

- Examining the accuracy of a predictive analytics tool begins by observing the distribution and outcomes by risk score, and if scores are translated into classifications, observed outcomes by risk classification. This most simple look at the data can illustrate how meaningful the distribution is and if outcome prevalence increases with a corresponding increase in risk classifications.



- Examining equity begins with examining outcomes by risk classification within and across racial/ethnic groups. For a tool to be equitable, the classifications of low, moderate, and high risk, for example, must mean the same thing regardless of race/ethnicity. In other words, families classified as high risk within each group must have similar outcome rates. Thresholds may need to be shifted across groups as well as within specific groups to achieve this.
- Statistical methods exist to help achieve equity in results; see Gottfredson method (Gottfredson & Snyder, 2005) to capture residual. Silver, Smith, and Banks (2000) suggest multiple-methods approaches. Bivariate correlations and regression analyses should be examined in so far as they can inform the relationship between risk factors, outcomes, and critical population subgroups.
- For example, the Receiver Operating Curve (ROC) is a measure of accuracy that represents the range of sensitivities and specificities for a test score. A test with perfect identification of positive cases is represented by the upper-left corner of the plot (100% sensitivity and 100% specificity), whereas a test with correct identification at random would result in a 45-degree diagonal from the lower left to the upper right.

- The closer an ROC curve is to the upper left corner, the more accurate the test. The area under the curve (AUC) statistic depicts how well the test separates the group being tested into those with and without the negative outcome in question. Many statistical tools are available to understand how a model is operating. Researchers should examine the model results using all test statistics to develop the best possible understanding of the model.

Refining Models

Predictive models are sensitive to their data inputs. As populations shift, new policies are implemented, or data quality changes, predictive models need to be refined. Therefore, all predictive models should be supported by an ongoing continuous quality improvement effort. Outcomes should be continuously analyzed and examined against a criterion for accuracy and equity among subgroups. If criteria are not met, refinements such as scoring

All predictive models should be supported by an ongoing continuous quality improvement effort.

threshold adjustments, new variable inclusion, or variable deletion can be made as long as they are done within guidelines for responsible use of predictive modeling.

Refinements must be made with the same statistical rigor used to create and analyze the original predictive model. Additionally, information collected from testing, piloting, or implementation may hold the potential to better understand and refine the model. These data and all other information should be considered when refining a model.

If the analysis reveals that the model cannot be refined to function validly, revisit the plan with relevant stakeholders to explore possible next steps, which may include, but are not limited to (1) prospectively collecting additional risk and protective factors under the hypothesis that accurate identification of additional, previously not measured items may improve equity of the model; or (2) developing a consensus and evidence-informed decision-support approach to structuring the decision being examined.

Core Methods for Implementation and Evaluation

Evaluating a model's use and fidelity to implementation is critical to understanding its impact and can involve both qualitative and quantitative measures. Evaluating impact relative to intent is best accomplished with a rigorous evaluation design, such as quasi-experimental or randomized control trial design. Essentially, commit to testing assumptions about the "intent versus impact" of predictive model use during pilot and implementation. This can include:

- Qualitative methods (through coaching, case reviews, and staff/case supervision); and
- Quantitative methods such as monitoring workers' resulting risk classifications relative to case decisions and extending those methods to rigorous evaluation of interventions.

Implementation can be supported by the following actions.

- Allow staff to override the policy/practice recommendation driven by the predictive model/tool.
- Monitor overrides and other aspects of practice (CQI) to support implementation and practice integration and enable process evaluation to assess fidelity.
- Integrate strategies from implementation science and innovation theory such as piloting



with an early adopters group to test and refine models and their application, then use this group to disseminate successes and onboard the early majority group.

- Develop and adhere to a rigorous communication plan and strategy, particularly with agency and community stakeholders, to increase transparency and build trust.
- Identify and train to the practice skills that front-line staff and agency decision makers need to both strengthen model inputs and apply model outputs. For example, structuring decisions with predictive analytics to improve practice will be more effective if implemented by line staff with family engagement and assessment skills, supported by skilled supervisors and managers with sufficient resources.
- Explore methods to test utility and fidelity in a pilot or other implementation stage such as (1) observing how workers use and record information; (2) creating a feedback loop; (3) conducting confidential surveys and/or focus groups; (4) observing, if possible, fidelity and impact on outcomes; and (5) developing the process for evaluating intent versus impact.

In Conclusion



“Big data” and predictive analytics approaches hold tantalizing opportunities for advancing evidence-based, data-driven practice and policy. Given how ubiquitous these technologies and methods already are in our daily lives, the question before us is not “if” but “when,” which makes the question of “how” an urgent one.

This paper calls out issues we must examine as we consider how we might use predictive analytic models to inform decisions about families’ lives.

We owe it to children and families to question claims about what these models can do, as the responsibility for the decisions ultimately lies with us, the human beings who must carry them out.

Whether we are researchers, agency administrators, legislators, advocates, social workers, or families involved in the system, we must demand and commit to an ethical standard that is understood, shared, and endorsed by all.