ASPE ASSISTANT SECRETARY FOR PLANNING AND EVALUATION
OFFICE OF HUMAN SERVICES POLICY

REPORT
September 29, 2022

Avoiding Racial Bias in Child Welfare Agencies' Use of Predictive Risk Modeling

Actions child welfare agencies can take to mitigate the likelihood that predictive risk models will exacerbate racial and ethnic disparities include carefully examining the data used to develop these models, evaluating model performance and impact, engaging community members throughout the process, and achieving transparency about how the models are used.

Danielle Whicher, Emma Pendl-Robinson, Kyla Jones, and Allon Kalisher

KEY POINTS

- Several child welfare agencies in the United States have implemented or are developing predictive risk models (PRMs) to help inform child welfare decisions.
- PRMs can assist caseworkers by identifying previously unobserved patterns in data related to future outcomes. However, PRMs also have the potential to exacerbate existing racial and ethnic disparities if not developed with careful attention to the methods used.
- Child welfare agencies and their vendors can take various actions to identify and mitigate racial and ethnic bias when planning, developing, implementing, and monitoring a PRM.
- When planning to develop a PRM, agencies should identify an appropriate use case and available data sources to use in supporting the PRM development.
- When developing a PRM, agencies and vendors should assess the quality of the available data sources, make careful decisions about PRM parameters, validate the outcome variable they plan to predict, and assess its predictive performance for important subgroups.
- When implementing a PRM, agencies should assess the risk and potential benefits of using the model, consider how best to balance PRM misclassification rates, and determine how to present predictions to agency staff and train them in using those predictions.
- When monitoring a PRM that has been implemented, agencies should continuously examine its predictive performance and evaluate its impact on equity and other outcomes.
- In each phase, agencies should engage community members and caseworkers in making key decisions and be transparent about their approach to developing and using PRMs.
- Federal agencies could consider supporting the integration of equitable, high-quality PRMs by (1) refusing to fund the development of proprietary PRMs; (2) developing guidelines on approaches for engaging with community members, governing the use of PRMs, and the public release of information about how PRMs are being used; and (3) working with child welfare agencies to conduct rigorous evaluations of these models.

INTRODUCTION

Overview of Child Welfare in the United States

In 2020, about 3.9 million referrals involving about seven million children were made to child protective services (CPS) across the United States (U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau, 2022a). That same year, there were just over 400,000 children in foster care (U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau, 2022b). Overall, about 37% of all children in the United States experience a CPS investigation by the age of 18. Rates of child welfare investigations are greatest for certain racial and ethnic groups, with African American and Hispanic children experiencing an investigation by the age of 18 at rates of approximately 53% and 32%, respectively, compared to 28% of White children (Kim et al., 2017). Additionally, children from certain racial and ethnic groups are overrepresented in the foster care system. For instance, while American Indian and Alaska Native children and African American children made up just 1% and 14% of the U.S. child population in 2019, respectively, they accounted for 2% and 23% of the foster care population (Child Welfare Information Gateway, 2021).

Child welfare agencies receiving referrals for child maltreatment must make a number of decisions about them, as well as the families already involved in child welfare services. These decisions include the following: (1) screening decisions based on referral calls, including deciding whether an investigation is warranted and triaging investigations based on child safety or maltreatment risk classification; (2) decisions about removing children from their families based on the results of those investigations; (3) decisions about moving children from one foster care placement to another; (4) decisions about returning children to their families; (5) assessments of risk of needing foster care or child welfare services; and (6) decisions about allocating resources and interventions across communities (Vaithianathan, Putnam-Hornstein et al., 2019). The decisions that child welfare agencies make can have a significant impact on children, families, and communities; in addition, child welfare decision-making processes can vary substantially across states and regions (Casey Family Programs, 2018; Damman et al., 2020).

Although evidence shows that families from certain racial and ethnic groups are overrepresented in their involvement with CPS (U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau, 2022a), there are outstanding questions related to the extent to which this pattern is the result of biased decision-making on the part of child welfare agencies. Researchers have developed six main explanatory pathways for this disproportionality: (1) disproportionate and disparate needs of children of different racial and ethnic backgrounds; (2) racial bias and discrimination by individuals, such as protective services caseworkers and mandated reporters; (3) child welfare system factors, such as lack of resources for families of different racial and ethnic backgrounds; (4) geographic context; (5) policy and legislation; and (6) structural racism (Child Welfare Information Gateway, 2021). However, there is a lack of evidence demonstrating the influence various social factors have had on this overrepresentation compared to the influence of the decision-making processes of individual child welfare agencies.

Evolution of Decision-Making Processes in Child Welfare Settings

Child welfare agencies have traditionally relied heavily on human decision making alone to make determinations about the risk of child maltreatment and adverse outcomes (Cuccaro-Alamin et al., 2017). However, given the large number of referrals and limited resources, these agencies can experience challenges identifying and protecting children at risk (Vaithianathan, Putnam-Hornstein et al., 2019). As a result, over the last three decades, agencies have adopted various tools designed to assist staff in making informed and efficient decisions that ultimately reduce the risks of negative outcomes for children and families.

In the early 1990s, child welfare agencies began using consensus-based and actuarial tools to inform decisions (Baird et al., 1999). Consensus-based tools rely on expert theories on the factors that lead to child maltreatment (Cuccaro-Alamin et al., 2017; Drake et al., 2020). These tools then were superseded by actuarial tools, which empirically identify, weight, and combine risk factors related to abuse and neglect to classify a family's or individual's risk of these outcomes (Shlonsky and Wagner, 2005). Research on the use of actuarial tools has shown that those child welfare agencies using them achieve "greater predictive validity and interrater reliability" than agencies not aided by these tools, leading to their wide adoption over the last 20 years (Cuccaro-Alamin et al., 2017; Drake et al., 2020). A limitation of actuarial tools is that variables included in them are those that research has empirically demonstrated to have a relationship to child maltreatment. Therefore, actuarial tools could be missing factors that are predictive of child maltreatment and abuse but do not have a demonstrated relationship to these outcomes (Cuccaro-Alamin et al., 2017; Drake et al., 2020).

More recently, a few researchers and child welfare agencies have begun developing predictive risk models (PRMs) to support child welfare decisions. Lanier et al. (2020) define predictive analytics as a sophisticated form of risk modeling that uses historical data to understand relationships between myriad factors to estimate a probability score for the behavior or outcome of interest. This form of analytics uses machine learning or other data processing techniques to provide information for "decisions, judgments, and/or policy implementations that impact opportunities, access, liberties, rights, and/or safety" (Pittsburgh Task Force on Public Algorithms, 2022). PRMs are intended to assist caseworkers in synthesizing data from various sources to inform decision making (Cheng et al., 2022). However, given the overrepresentation of African American, Hispanic, and American Indian and Alaska Native children in the child welfare system, some have raised concerns that such models may inadvertently incorporate racial and ethnic biases, which could ultimately further increase existing disparities and degrade trust in CPS.

Although PRMs have a variety of potential applications in child welfare, several of the relatively small number of tools currently in use or under development are designed to inform front-end decisions, such as those about the need to investigate referrals made to child welfare hotlines (Exhibit A). This design choice likely has been made because at this initial stage, before any interaction with families occurs, case workers have limited clinical information to inform their decisions but there is a great deal of structured data that a PRM can incorporate to predict what might happen if a particular referral is not investigated. PRMs have also been or are being developed to inform policy decisions related to allocating resources effectively. For example, the Administration for Children's Services in New York City created a PRM to identify risk factors for families who had "frequent involvement" with CPS. It was used for macro-level planning for the child welfare agency's resource allocation (Chapin Hall and Chadwick Center, 2018).

Exhibit A. Examples of PRMs Used in Front-End Screening Decisions

Allegheny Family Screening Tool (AFST) is used by the Allegheny County Department of Human Services in Allegheny County, Pennsylvania. The PRM advises child welfare hotline staff to triage children based on the likelihood of future involvement with the child protection system (Vaithianathan, Putnam-Hornstein et al., 2019).

Douglas County Decision Aide (DCDA), used by the Douglas County, Colorado Department of Human Services, is a hotline screening tool that predicts the risk of a child being removed by CPS to determine whether an investigation should be opened (Fitzpatrick & Wildeman, 2021).

Safety at Screening, implemented by the Oregon Department of Human Services, was a real-time decision support tool to provide accurate and equitable indications of the likelihood of future adverse outcomes for children named in reports of abuse and neglect. This PRM was phased out in June 2022 (Oregon DHS, 2019).

Report Approach and Objectives

This report describes best practices aimed at preventing and mitigating racial and ethnic bias in child welfare agencies' use of PRMs. To identify best practices, we first completed an environmental scan of relevant literature and other resources that describe the potential pros and cons of using PRMs in child welfare, and approaches for addressing racial and ethical bias in PRMs. As part of this scan, we also identified relevant articles from the criminal justice and health care literature to consider what child welfare could learn from these other fields, which have more experience with these types of models. We confirmed that we had identified the most relevant resources by speaking with five key informants. Key informants included individuals with experience developing PRMs for child welfare agencies as well as individuals with expertise considering the ethical issues related to the use of these models.

We developed a structured outline after reviewing the resources identified through the environmental scan. We shared this outline with a broader group of experts and hosted two virtual roundtable discussions to receive feedback on its content. Roundtable participants included representatives from child welfare agencies that have implemented or considered implementing PRMs; researchers who have developed PRMs for child welfare settings; advocates experienced in thinking about the issues of using PRMs in a child welfare setting; and scholars who have considered the ethical implications of using these models to inform child welfare decisions. During the roundtable, we asked about information that might be missing from the outline or points in it with which participants disagreed. We revised the report outline based on the feedback we received.

The resulting report has four main objectives:

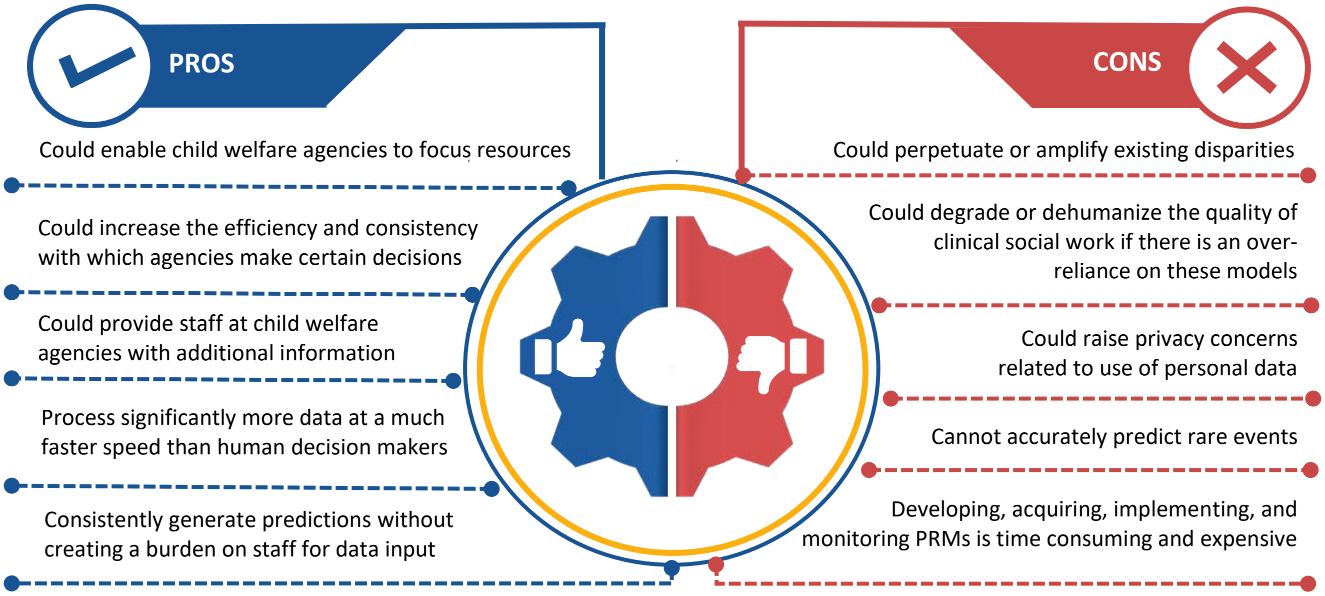
1. To describe the potential pros and cons of using PRMs in child welfare settings
2. To review actions developers and child welfare agencies can take to detect and mitigate risks of racial and ethnic bias in these models
3. To discuss the significance of transparency and explainability in promoting trust in PRMs, including what information about a model should be disclosed to people using and affected by it
4. To identify steps federal agencies can take to promote fairness in the use of PRMs for child welfare

PROS AND CONS OF USING PRMS TO INFORM CHILD WELFARE DECISIONS

Although PRMs have the potential to improve decisions child welfare agencies make, there are also some risks of harm associated with using these models. The specific pros and cons of a particular PRM use case will vary. In this report, we describe the broad types of potential pros and cons that may be associated with various PRM applications when compared to current decision-making processes, which are typically informed by policy or practice frameworks, sometimes with the support of actuarial models and other tools (Exhibit B).

Given the significant impact that child welfare decisions can have on children and families, agencies considering implementing a new PRM should carefully review these pros and cons, and engage in discussions with researchers, community members, advocates, and agency staff to understand their perspective regarding these issues.

Exhibit B. Potential Pros and Cons of Using PRMs to Inform Child Welfare Decisions



PROS

- Could enable child welfare agencies to focus resources
- Could increase the efficiency and consistency with which agencies make certain decisions
- Could provide staff at child welfare agencies with additional information
- Process significantly more data at a much faster speed than human decision makers
- Consistently generate predictions without creating a burden on staff for data input

CONS

- Could perpetuate or amplify existing disparities
- Could degrade or dehumanize the quality of clinical social work if there is an over-reliance on these models
- Could raise privacy concerns related to use of personal data
- Cannot accurately predict rare events
- Developing, acquiring, implementing, and monitoring PRMs is time consuming and expensive

Potential Pros of PRMs

When used to assist in child welfare decision making, PRMs have the potential to positively influence a number of outcomes. A PRM could enable a child welfare agency to shift its approach from reactive decision making to proactively providing communities and families with resources that could help prevent negative outcomes. For example, a PRM could support a child welfare agency's effort to identify gaps in resource allocation or identify specific children and families who would benefit from additional resources (Lanier et al., 2020). In this way, PRMs can support the shift from a more punitive child welfare approach to one informed by a public health perspective, thus allowing agencies to use their resources for cases that require the greatest attention (Capatosto, 2017; Drake & Jonson-Reid, 2018). PRMs can shape conversations around different policy choices, potentially allowing child welfare agencies to be more responsive to families' needs (Chapin Hall and Chadwick Center, 2018).

PRMs can also provide agency staff with information they otherwise would not have because they can identify previously unobserved patterns in data that relate to future outcomes for children and families (Pryce et al., 2018). By providing this information, PRMs can increase the efficiency and consistency with which agencies make certain decisions, such as how to triage incoming cases and whether calls to hotlines require further investigation (Cuccaro-Alamin et al., 2017). Consistent decision making, informed by PRMs, can reduce disparities if some child welfare staff are making biased decisions in the absence of the model (Drake & Jonson-Reid, 2018). However, consistent decisions based on PRMs that reflect systematic racial or ethnic biases could have the opposite effect.

During our roundtable discussions, participants noted that there is less risk involved when using PRMs to inform positive decisions, such as those about whether to provide additional services to families, rather than negative ones, such as the decision to place a child in foster care. Participants attributed this to the fact that negative decisions, such as removal, have a greater direct impact on the family. A central goal of child welfare agencies is to ensure that the decisions staff make are fair and reflect actual risk; although not a replacement for staff experience and expertise, PRMs can provide supplemental information staff can use to inform these decisions.

Another advantage of PRMs is their ability to generate predictions without having to rely on new staff data entry, as is generally required for models developed using actuarial methods, such as structured decision-making (SDM). The increased reliance on human input for generating risk scores makes SDM and other models developed using actuarial methods more vulnerable to human error and biased coding (Chouldechova et al., 2018). As Drake et al. note, if a PRM is likely less subject to bias than current decision-making processes, it is morally defensible for child welfare agencies to use that model to inform staff decisions (2020).

Potential Cons of PRMs

PRMs have the potential to improve decision-making processes but may also be associated with an increased risk of negative outcomes. Researchers have noted various ways in which historically biased or low-quality data can negatively impact the performance of PRMs and, as a result, the children and families affected by those models. As discussed further below, if the data used to train and validate a PRM reflect child welfare staff's biased decision making, the same biases can be reflected in the PRM itself, which ultimately can lead the PRM to perpetuate or even exacerbate existing disparities (Cuccaro-Alamin et al., 2017; Drake et al., 2020; Feng & Wu, 2019; Samant et al., 2021; Yen & Hung, 2021). A related factor is that some government administrative data include more information on certain racial or ethnic groups compared to others because those groups are more likely to be involved in government programs (Chouldechova et al., 2018; Drake et al., 2020). If these data are used to create PRMs, those models could potentially create feedback loops in which certain racial and ethnic groups are flagged as higher risk and more likely to be investigated simply because there are more data about them (Chouldechova et al., 2018; Favaretto et al., 2019).

> *"Historically over-regulated and over-separated communities may get caught in a feedback loop that quickly magnifies the biases in these systems. Even with fancy—and expensive—predictive analytics, the family regulation system risks surveilling certain communities simply because they have surveilled people like them before."*
>
> -Samant et al. (2021)

A final consideration related to the data used to train and validate PRMs is that the definition of "neglect" has not been modified to reflect a differing understanding of the challenges related to caring for children in poverty versus serious issues of maltreatment by caregivers (Samant et al., 2021). This distinction is important because the U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau (2021) reported that 61% of all cases of substantiated child maltreatment involved only neglect. As a result, much of the current administrative data reflects social issues and racial differences associated with neglect due to poverty in the United States, rather than patterns of other forms of maltreatment. PRMs developed using these data will produce predictions that may perpetuate these patterns, effectively treating cases of neglect as a result of poverty and other serious issues of maltreatment by caregivers the same. A criticism of the PRM developed in one county, for example, described the algorithm as a tool that simply reports on how many public resources a family uses (Francis et al., 2022). It is critical to examine the present and historical realities of racial and ethnic bias in the child welfare space when selecting data to develop and validate these types of models.

In addition to the potential issues created by biased or low-quality data, child welfare agencies need to consider the resources required to develop and use a PRM, as well as the impact on staff decision making. The time and expense associated with PRM development and use should be carefully compared to the potential benefits of the model. The resources include those for developing and validating the PRM, as well as those needed to train staff, develop implementation plans and policy changes, inform the community about the use of the PRM, and evaluate the model's performance following implementation. If a PRM is unlikely to result in positive outcomes for children, families, and communities, resources may be better spent on alternative endeavors (Drake et al., 2020). Additionally, when agency staff place too much confidence in the outcomes of PRMs, these models can exert a significant degree of influence on decision-making processes that are supposed to incorporate human judgment, and in turn may degrade or dehumanize the quality of clinical social work (Capatosto, 2017; Drake et al., 2020; Drake & Jonson-Reid, 2018; Pittsburgh Task Force on Public Algorithms, 2022; Zytek et al., 2021).

Another potential issue with PRMs is that these models can raise privacy concerns related to the use of personal data and the failure to get consent (Cuccaro-Alamin et al., 2017; Drake et al., 2020). Although these data might be protected by normal government requirements, when community members are not informed or consulted regarding how their data are used to develop PRMs or otherwise inform child welfare decisions, the risk of public backlash and distrust regarding privacy concerns is increased. In addition, PRMs can risk causing increased distrust in CPS if they are used to make determinations that community members cannot challenge because they are not engaged in the development of the models or are unaware of how the models are being used (Francis et al., 2022).

A final limitation of PRMs is their inability to accurately predict rare events (Cuccaro-Alamin et al., 2017; Reisman et al., 2018). Although this consideration is important, roundtable participants noted that this limitation is not unique to PRMs; rather, it is shared by many tools, including the actuarial models that many child welfare agencies use. To be most effective, PRMs and other tools are generally designed to predict more common events, such as future child welfare system involvement, which may or may not be closely linked to rare events, such as maltreatment or death. As discussed further in the section on mitigating bias when developing a PRM, when more common proxy outcomes are used in these models, it is important for child welfare agencies to ensure that those outcomes are closely associated with rare events of greatest importance.

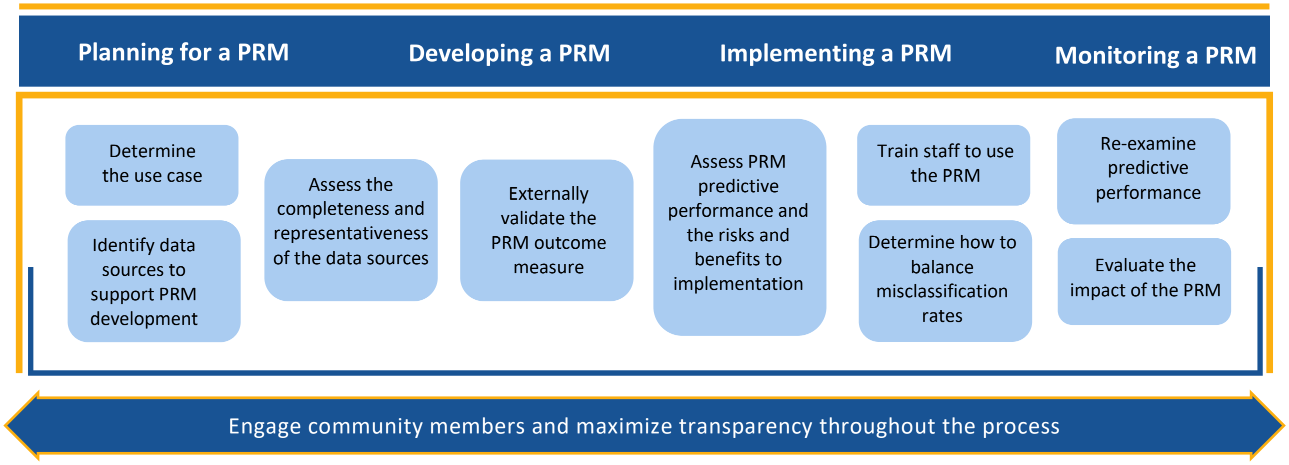
ACTIONS FOR IDENTIFYING AND MITIGATING THE RISKS OF RACIAL AND ETHNIC BIAS IN PRMS

Although there are some risks of harm associated with implementing PRMs in child welfare, should agencies and communities choose to use them, they can be reduced through a number of actions child welfare agencies can take during the process of integrating a new PRM. In this section, we describe the actions that child welfare agencies and PRM developers or vendors can take to reduce the risk that PRMs will perpetuate racial and ethnic bias or otherwise negatively impact families and communities. We categorize these actions into the following four groups that align with the different phases of integrating a new PRM into a child welfare agency's decision-making process (Exhibit C):

1. Planning for the development of a new PRM
2. Developing the PRM
3. Implementing the PRM within a child welfare agency
4. Monitoring and evaluating the performance of a PRM following implementation

The actions we describe for each phase are not linear, and although community involvement and transparency are highlighted at various points, child welfare agencies should involve community members, advocates, and child welfare staff and be transparent about their actions throughout all phases. Additionally, we recognize that the extent to which a child welfare agency and its partners can undertake each of these actions may vary depending on their available time and resources. However, if an agency is unable to undertake a specific action, it should be clear about the implications of not performing that action for the quality of their PRM and, more important, the likelihood that the PRM will lead to worse outcomes for children and families from various racial and ethnic groups.

Exhibit C. Actions for Addressing Bias in PRMs Designed to Inform Child Welfare Decisions


Planning for a PRM
Developing a PRM
Implementing a PRM
Monitoring a PRM
Determine the use case
Identify data sources to support PRM development
Assess the completeness and representativeness of the data sources
Externally validate the PRM outcome measure
Assess PRM predictive performance and the risks and benefits to implementation
Train staff to use the PRM
Determine how to balance misclassification rates
Re-examine predictive performance
Evaluate the impact of the PRM
Engage community members and maximize transparency throughout the process

Planning for the Development of a PRM

When planning for the development of a new PRM, agencies should thoughtfully identify the problem to be solved, the ultimate goal of the model, and the data sources available to support its development. During this process, actions that child welfare agencies can take to reduce the risk that PRMs under consideration perpetuate racial and ethnic biases include the following:

- Thoughtfully considering the goal of a PRM and how it will be used to inform agency decisions
- Engaging community members, advocates, regulators, and child welfare professionals in discussions about the PRM use case and other supporting decisions
- Evaluating the various data sources used to develop the PRM
- Ensuring transparency by contracting with vendors committed to sharing information about PRMs

Defining the goal of the PRM and how it will be used: When defining a goal for a new PRM, agencies should first consider a set of guiding principles to inform their work. These principles could include an explicit focus on reducing bias or child welfare system involvement in cases where it is not necessary to improve child well-being (Allen et al., 2020; The Annie E. Casey Foundation, 2020). Agencies should also assess the nature of their role and how implementation of a PRM will help improve their performance relative to that role. Based on this assessment, child welfare agencies can then define the specific problem they would like a PRM to help solve and consider whether it is the most appropriate tool for helping to address that problem. When evaluating whether using a PRM is appropriate for addressing a problem, roundtable participants suggested that agency leadership consider whether there is available and time-stamped data that could be used to develop a PRM and whether there is a distinct decision point the model could be used to inform. If there is not a distinct decision point, such as whether to investigate a case or whether to provide more resources to a specific community, a PRM is unlikely to be an appropriate tool for addressing an identified problem.

According to key informants with whom we spoke, once child welfare agencies have defined a goal and problem, they should consider how to position the information gathered from a PRM so it reaches decision makers at critical points that can positively affect the trajectory of a case. Specific questions that agencies can ask regarding how to position the information from a PRM include the following:

- Where should the information be delivered?
- When should it be delivered?
- How should it be delivered to ensure it is used to *inform* decisions, not *make* them?

- What are defining characteristics of cases that require action on the part of child welfare agency staff and what is an effective response to those cases?

Asking these questions during the planning phase can help child welfare agencies determine whether it is possible to use the information from a PRM to help improve decision-making processes. When considering how to deliver information from a PRM, agencies should aim to ensure that the information is considered along with other relevant information about a case. If the information from a PRM is delivered at a time when other information is not available, it is more likely that child welfare staff will use the information from the PRM to make a final decision about a case instead of considering it along with other relevant contextual factors. Applying these questions can also assist in more clearly defining the intended outcomes of a PRM. For example, the questions can help agencies make more uniform determinations about how families are deemed to be “high risk” and what interventions are appropriate in high-risk situations. Considering the use case before the PRM is developed can help state and local agencies develop a tailored approach to mitigate bias when they apply predictive modeling.

Engage community members, advocates, regulators, and child welfare professionals: Community members, advocates, child welfare professionals, and other end users should be given the opportunity to provide feedback on the PRM use cases developed, including the goals of the PRM, the problem it is designed to solve, and how it will be used (Cheng et al., 2022; Favaretto et al., 2019; Krakouer et al., 2021; Pittsburgh Task Force on Public Algorithms, 2022). It is also important to engage these groups in determining what definition of fairness the child welfare agency should use to interrogate the PRM, as stakeholders are likely to have varying perceptions of what fairness requires (Cheng et al., 2022). The selected definition of fairness should then inform the PRM development process, including what performance metrics agencies use to validate a PRM, as is discussed further in the section below on implementing PRMs (Chapin Hall and Chadwick Center, 2018; Purdy & Glass, 2020). Involving community members from the outset to ensure that the trade-offs associated with developing and implementing a PRM align with community values is important in building trust.

Agencies and researchers have adopted various approaches for engaging community members and other end users in the planning process. One approach is to establish an ethical review panel with broad representation of interested parties and those affected by a PRM (Chapin Hall and Chadwick Center, 2018). In addition to the use case, the review panel can also provide feedback on the development of the PRM and its implementation (Chapin Hall and Chadwick Center, 2018). Another approach is to co-design a code of ethics with community members, advocates, and child welfare professionals (Exhibit D) (Cheng et al., 2022). This code of ethics can then govern the identification of an appropriate PRM use case, its goals, and plans for its implementation. Such examples show how the community can be continually engaged, rather than being ancillary to the planning process. Types of decisions in which community members and end users can be engaged during later stages of the PRM planning process are described further in subsequent sections.

Consider the strengths and weaknesses of various data sources that could be used to train and validate a PRM: During the planning phase, agencies should carefully consider what data sources are available to support the training and validation of the PRM, because as noted previously, the data used can have a significant impact on the types of racial and ethnic biases that may be reflected in a model. Administrators should evaluate the strengths and weaknesses of various data sources, including historical and human cognitive biases that could be reflected in the data, and the data quality (Chapin Hall and Chadwick Center, 2018; Chouldechova et al., 2018; Lee et al., 2019; Pittsburgh Task Force on Public Algorithms, 2022).

Roundtable participants suggested that child welfare agencies should first consider how historical or human biases might be reflected in a PRM's outcome measures, because these measures will have a more significant impact on the performance of a PRM than the predictor variables included in the model.

Exhibit D. Engaging community members and end users to develop guiding principles for developing PRMs: a review of Douglas and Larimer Counties

In 2017, Douglas County's Child Welfare Division and Larimer County's Division of Family, Youth and Child Services partnered with the Centre for Social Data Analytics and the University of Auckland to develop a child welfare predictive risk model (Vaithianathan, Dinh et al., 2019). The Douglas County Decision Aid (DCDA) and the Larimer Decision Aid Tool (LDAT) were developed to support child welfare personnel with screening and triaging cases of alleged child maltreatment and neglect (Vaithianathan, Dinh et al., 2019).

The counties also joined with community members, end users, and other private partners to develop a practice profile. The purpose of the profile was to operationalize the guiding principles and core practices of their approach, making it repeatable for other counties (Metz, 2016). As a part of these efforts, each county sought to modify its current and existing policies to build an infrastructure in which the use of PRMs can supplement rather than replace clinical decision making.

The guiding principles identified in the practice profile represent various best practices identified in the literature, including due consideration for equity concerns arising from biased data sources, the centering of clinical judgment, a commitment to reduce biased decision making, and a commitment to transparency. The core practices define specific stakeholder activities, including communication with community members through the development of an advisory panel, evaluation of machine learning and other decision aids, and dedicated funding for the development of policies and protocols to manage the tools and train staff.

Though aspirational, the core practices and guiding principles can serve as an example of the real-life infrastructure needed to support child welfare staff and mitigate bias when planning for the development of PRMs.

When considering the potential for historical or human biases to be reflected in data, agencies should determine whether there is any evidence to suggest that decisions made by child welfare staff or other agencies are biased, and how that bias might affect the available data (Lee et al., 2019). For example, if a child welfare agency is using data drawn from staff data entry historically, it should consider the accuracy and completeness of that data. Additionally, if a child welfare agency is considering using arrest data in a PRM, it is important to consider that such data could be biased due to disparities in policing practices across neighborhoods (Pittsburgh Task Force on Public Algorithms, 2022). Agencies should also assess whether they are likely to have more information on certain subgroups of the population due to the involvement of those subgroups in social programs. If there are more data about certain subgroups, agencies should consider how this could impact the performance of the PRM once it is developed.

Key informants with whom we spoke also noted that agencies should consider when the data were captured and when they gain access to those data. If there were significant policy or practice changes between the time the data were captured and when an agency has access to them, such data may not be useful for a PRM. Additionally, the age of the data could be used to determine how certain variables are coded. For example, if an agency is planning to use arrest data dating back many years, it should consider creating a variable that captures time since the most recent arrest. That variable may be more predictive of future child maltreatment than a dichotomous variable of whether a person has a prior arrest or not. If weaknesses or biases are identified in the data sources a child welfare agency is considering using to develop a PRM, agency staff and developers should work together to address those weaknesses or select alternative data sources, if feasible (Favaretto et al., 2019).

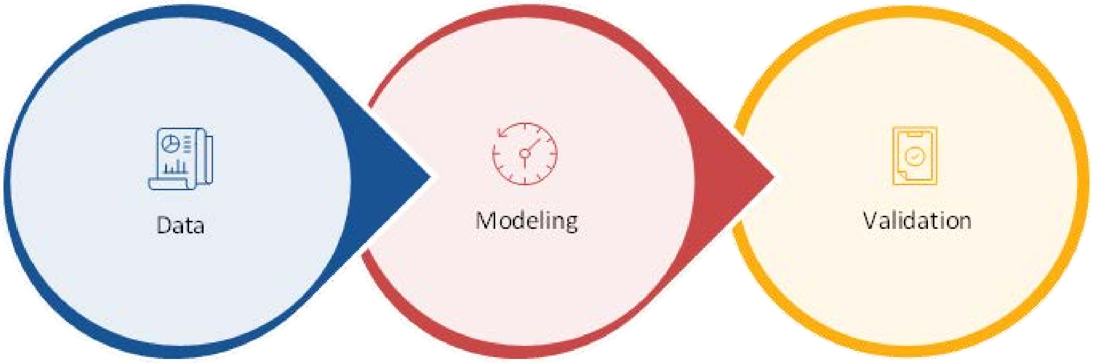
Contractually obligate all vendors to waive proprietary information: Child welfare agencies that rely on private vendors to develop their PRMs should ensure transparency by contractually obligating vendors to waive proprietary or trade secrecy information surrounding the PRM and its predictive performance (Reisman et al., 2018). Trade secrecy surrounding the development and validation of PRMs can discourage meaningful public participation and government or agency accountability (Francis et al., 2022). As discussed further in the section on transparency and explainability, the types of information that vendors should be required to share with the child welfare agency and the public include the data and methodology used to develop the PRM, the predictor and outcome variables included in it, and the predictive performance of the model overall and for specific subgroups.

Developing a PRM

After planning for the PRM, the child welfare agency or its vendor builds the model (Exhibit E). During this stage, agencies can take numerous actions to identify and mitigate the risks of racial and ethnic bias, including the following:

- Assessing the representativeness and completeness of the training and validation data
- Ensuring that subject matter experts in child welfare and PRMs work together to make informed decisions about how to specify the PRM parameters
- Externally validating the PRM outcome variable against historical data
- Assessing the predictive performance of PRMs overall and for different racial and ethnic subgroups
- Continuing to engage program staff, community members, and others through case review exercises to seek feedback on the development process and PRM design

Exhibit E. Overview of PRM Development Process



Assessing the completeness and representativeness of training and validation data: When building a PRM, child welfare agencies and the developers or researchers with which they are working should assess the completeness and representativeness of the data sets they plan to use, including the equitable representation of racial and ethnic subgroups. If a significant amount of data are missing from a data set an agency is considering using, the agency should consider the reasons those data are likely missing and how that may affect the modeling process. In cases where the missing data are likely to introduce bias into the modeling process, agencies should select alternative data sources. Additionally, over- or underrepresentation of subsets of the population in the training data set can affect the generalizability of the model (Cahan et al., 2019; Lee et al., 2019). If the training and validation data are not representative of the population to which the PRM will be applied, developers should consider reweighting variables. For example, if the individuals represented in the training data are, on average, older than the population of interest, developers should reweight the age variable.

Constructing PRM parameters: Child welfare agencies should ensure that subject matter experts in child welfare and experts in developing PRMs work together to make informed decisions about how to construct these models. Decisions that have important ethical implications are made throughout the model development process (Barocas & Boyd, 2017). They include, for example, decisions about creating a test data set, the choice of a learning algorithm, and acceptable model error rates (Barocas & Boyd, 2017). In addition, child welfare agencies must define the outcome of interest, which involves thinking through common issues such as the appropriate outcome windows and the temporal availability of the outcomes of interest. Bias can be introduced into the modeling process during any of these steps depending on these decisions and issues. Having a diverse team of experts and consulting community members and other end users when making choices about the modeling approach can help child welfare agencies ensure that the decisions they make are well justified and publicly defensible. During this process, child welfare agencies and model developers should be thoughtful about minimizing the effects of race and ethnicity in future decision making, understanding that demographic predictors should be incorporated and interpreted cautiously in the overall modeling approach (Chapin Hall and Chadwick Center, 2018; Chouldechova et al., 2018; Kennedy et al., 2021).

Externally validating the PRM outcome variable: Child welfare agencies are concerned about whether children are at risk of negative outcomes, such as maltreatment or death. However, as noted previously, predicting outcomes such as maltreatment or death can be statistically challenging because these outcomes are rare events (Lanier et al., 2020). Therefore, most PRMs used in the child welfare setting predict more common outcomes related to future involvement with the child welfare system, and there is a risk that these proxy outcomes may not be closely correlated with the primary, ground truth outcome of interest.

To determine how closely correlated proxy outcomes are with “ground truth” outcomes of interest, child welfare agencies should validate the proxy outcomes against other external indicators related to maltreatment, such as hospitalization data from claims or death records (Chouldechova et al., 2018; Vaithianathan, Kulick et al., 2019). For example, the Allegheny Family Screening Tool (AFST) V2 was created to predict the likelihood of future child welfare out-of-home placement. Researchers wanted to validate that the PRM could predict external variables also associated with child maltreatment by using hospital data. The hypothesis was that children identified as being at higher risk of out-of-home placement would also be children with a higher risk of being hospitalized due to abuse-related injuries. Through the external validation analyses, the researchers found a positive correlation between the AFST V2 risk scores and hospitalization due to abuse-related injuries for analyzed subgroups, thus supporting the conclusion that the PRM outcome was associated with the primary outcome of interest (Vaithianathan, Kulick et al., 2019).

Assessing predictive performance across subgroups: To understand whether a PRM performs equally across subgroups, child welfare agencies should assess its predictive performance for the overall population and different racial and ethnic minority groups (Chouldechova et al., 2018; Chouldechova & G’Sell, 2017; Drake et al., 2020). One common approach used to compare performance across groups is examining the area under the curve (AUC) for those groups.[1] When AUCs differ between subgroups and the model does not perform well for a specific subgroup, developers should take steps to enhance its performance for that group without sacrificing overall model performance, or it should not be used (Chouldechova et al., 2018). However, in some situations, such as when a PRM is predicting a rare outcome or when false positives and false negatives from a PRM are weighed differently by child welfare agencies, AUCs are not the best approach for examining

> *“Even when two models have comparable overall performance, they may nevertheless disagree in their classifications on a considerable fraction of cases.”*
>
> – Chouldechova & G’Sell (2017)

[1] The AUC is used to compare the accuracy of predictions across subgroups or statistical models. AUC values can range from 0.5 to 1 with 0.5 representing no apparent accuracy and 1 representing perfect accuracy (Hanley & McNeil, 1982).

predictive performance. It is therefore important for child welfare agencies and researchers those agencies are working with to carefully consider their approach.

Child welfare agencies also should compare the predictive performance of newly developed PRMs overall and across subgroups to that of any tool or tools they are currently using. PRMs should be implemented only when they perform better than current tools (Drake et al., 2020; Drake & Jonson-Reid, 2018).

Engaging community members and model users in case review exercises: Child welfare agencies should continue to engage program staff, community members, and advocates to get feedback on how the PRM works and how it could be used to inform decisions. Participants in our roundtable discussions recommended completing case reviews throughout the modeling process with child welfare agency program staff who would be using the PRM. The case reviews are an effective way of explaining to a nontechnical audience how a PRM operates and soliciting feedback about the modeling approach. The case review exercises consist of walking through examples of the predictions or outcomes a PRM would provide for a specific case, such as a hotline call the agency might receive. After walking through a case, staff would discuss how they would handle the case and whether they agree with the PRM prediction. If staff disagree with the PRM outcomes, agency leaders and model developers can ask them why they disagree and whether they think there should be changes to the model. This iterative process of doing case reviews and incorporating feedback into the model builds trust between the staff and the PRM development team.

Child welfare agencies can also engage in case reviews with community members. Roundtable participants suggested that agencies provide additional context to community members about how decisions are currently made and how the PRM would inform those decisions if implemented. For example, if a PRM is replacing an actuarial model, child welfare agencies should explain how the existing actuarial model works and what the outcome of that model would have been for the example case compared to the outcome for the new PRM. If community members raise concerns about the outcomes of a PRM, child welfare agencies and developers should consider the best approach to address and mitigate those concerns. In cases where agencies and developers do not believe that changes are required to address a concern raised, they should explain why. Child welfare agencies should also consider engaging staff and community members in conversations about the variables or features the PRM will include, as well as thresholds for different risk levels. Roundtable participants suggested that if community members or staff are uncomfortable with certain features, such as race, ethnicity, or geographic indicators, the agency should consider whether these features should be removed from the model.

Implementing a PRM

Implementing a PRM involves making decisions about how the predictions from the tool should be used, preparing agency staff to use it, integrating it into existing information technology infrastructure, testing it to ensure it works correctly, and finally launching its use in practice. During this implementation process, actions that agencies can take to reduce the risk that the model will exacerbate racial or ethnic bias or bias in child welfare decision making include the following:

- Conducting an assessment to weigh the risks and potential benefits of implementing the PRM
- Considering how best to balance misclassification rates and mitigate the negative impacts of any potential errors
- Determining how best to present the predictions from a PRM to end users so they use the information in the way the agency intends
- Training staff so they understand the PRM and how the predictions should be used
- Ensuring that the PRM is explainable and its development and use are transparent

Conducting a risk and benefit assessment: Child welfare agencies should assess the risks and potential benefits of all new or modified PRMs they plan to implement in comparison to existing decision-making processes (Pittsburgh Task Force on Public Algorithms, 2022; Reisman et al., 2018). This review should consider the potential impact of a PRM on "fairness, justice, bias, or other concerns across affected communities" (Reisman et al., 2018) and weigh any potential benefits of using the model against potential or expected problems resulting from that use (Pittsburgh Task Force on Public Algorithms, 2022).

Agencies should follow several steps when conducting these assessments. The first step in this process requires that they clearly define how decisions are being made in the absence of the PRM and the risks and benefits of making decisions using that approach. For example, agencies should specify what policies and tools are in place to guide decision making, and the pros and cons of those policies and tools. This step might require directly asking staff, community members, and advocates about their opinions regarding these policies and tools. The second step is to define how the new or modified PRM will change current decision-making processes, including any impact that using the PRM will have on existing policies and tools, and the risks and potential benefits of making those changes. As with the previous step, it is important to confirm with end users and those affected by the model that the identified risks and benefits are accurate.

In both steps, agencies should consider both the "harms of allocation" that result in some subgroups of the community receiving fewer resources than other subgroups, and the "harms of representation" that result in a system reinforcing the differential treatment of some subgroups in comparison to others (Reisman et al., 2018). Agencies should also identify approaches for minimizing or mitigating any potential negative effects of a PRM on end users and affected communities. Finally, they should weigh the risks and benefits to ensure they are making informed decisions about the appropriateness of the new or modified PRM.

Determining how to balance misclassification rates: Child welfare agencies implementing PRMs will also have to consider how best to balance different types of prediction errors (Lee et al., 2019). For example, for PRMs that will be used to inform screening decisions, agencies will have to consider how to balance harms associated with investigating families less likely to require a child welfare intervention with those associated with not investigating a family for whom early intervention could have prevented future negative outcomes for a child. These considerations are important because both unnecessary intervention and failing to intervene when needed can have significant consequences. Unnecessary intervention by the child welfare system can cause psychological harm to families, whereas failing to intervene when necessary can result in a child experiencing additional maltreatment or—on rare occasions—even death (Cuccaro-Alamin et al., 2017; Drake & Jonson-Reid, 2018).

There are several approaches to estimating misclassification rates of PRMs. When determining which approach to use and how to balance different considerations related to misclassification rates, agencies should refer to the goals the PRM was designed to help achieve and the definition of fairness the agency has agreed on with end users, community members, and advocates (Lee et al., 2019; Rajkomar et al., 2018). The most common approach to examining error rates is looking at the sensitivity and specificity of the PRM overall, and for different racial and ethnic subgroups.[2] A PRM with a high sensitivity and low specificity threshold will help identify all families or communities that require intervention or additional resources but will also identify some families and communities that do not require a child welfare intervention. Conversely, a PRM with a low sensitivity and high specificity threshold will correctly screen out families and communities not requiring child welfare intervention or resources but will also fail to identify some families and communities for whom intervention or resources could be beneficial (Cuccaro-Alamin et al., 2017). Agencies must weigh these

[2] Sensitivity measures how often a PRM correctly assigns a higher probability to cases that have a higher risk for the outcome being predicted; specificity measures how often a PRM correctly assigns a lower probability to cases that have a lower risk for the outcome being predicted (Drake et al., 2021).

different considerations when creating prediction thresholds for PRMs they plan to implement (Drake et al., 2020; Park et al., 2021; Pryce et al., 2018). They may also want to consider alternative approaches to estimating misclassification rates, such as calibration curves or positive and negative predictive values.

Once an agency determines appropriate thresholds for recommended actions based on the predictions from a PRM, it should determine how to mitigate any potential adverse effects from the model. For example, if an agency implementing a new PRM to inform screening decisions decides to set the threshold for recommended child welfare agency contact so the threshold is highly sensitive but not very specific, it should consider how it will minimize the impact of unnecessary intervention on families for whom the risk of future maltreatment or neglect turns out to be low.

Determining how to present the predictions from a PRM to end users: The predictions or recommendations from a PRM can be presented to end users in a variety of formats, and the way such information is presented can affect how they perceive it. For example, a roundtable participant noted that if the prediction is represented with colors, such as red, yellow, and green, it is important to consider how end users understand those colors and how that method of presentation may impact their ultimate decision. The same is true for predictions presented as high, medium, or low risk; it is important to understand how end users perceive these different risk categories and how the information can influence their decision making.

To understand how the presentation of information affects the way in which end users incorporate that information into their decision making, child welfare agencies should engage the staff members who will use the information in individual or small group discussions, again potentially using case review exercises that present predictions in different formats. The case review discussions could focus on the ways in which the format of the prediction influences how end users use the information to inform their decisions and their likely decision about the case. Based on the feedback received during these discussions, child welfare agencies can select a format that aligns with their goals and should develop written policies, protocols, or procedures documenting how staff should use the information from the PRM to inform their decisions.

Developing and implementing a learning plan for agency staff on how to use a PRM: Once a child welfare agency determines how best to balance misclassification rates from a PRM and how to present the predictions to staff and other end users, it should develop a learning plan to train staff on the use of the PRM and provide opportunities for those staff to apply that training to ensure the PRM is used appropriately. Before implementing the learning plan, agency leaders should ensure they have a strong understanding of how the PRM was developed and its intended use. This knowledge includes, at a minimum, understanding the basics of how the model was developed, the modeling methods employed, the model outcomes or predictions, the model's accuracy, its benefits and limitation, and how and when the predictions should be used (Cuccaro-Alamin et al., 2017).

When training agency staff and other end users, in addition to conveying key information about the model development approach, model performance, and the goals of using the PRM, agency leaders should describe how information from the model will be integrated into existing systems and review those policies or protocols that document how staff should use that information when making decisions. Being clear with staff and other end users about the goals of a PRM and how its information should be used is particularly important because staff's understanding of these components will impact their ability to manage the tool and make decisions effectively (Cuccaro-Alamin et al., 2017; Elgin, 2018). In general, child welfare agencies should prioritize human judgment when integrating PRMs into service delivery or agency operations (Chapin Hall and Chadwick Center, 2018; Cheng et al., 2022). Therefore, in most cases, agency leaders should advise staff to use the information from the PRM to help inform their decisions but also emphasize that staff should review all of the other information they receive about a case.

Once the training is complete, the agency should provide opportunities for staff to apply the training to selected cases. The agency should also provide ongoing supervision or coaching of staff and other end users to ensure they are using the PRM as intended. Child welfare agencies might need to conduct repeated trainings if the following occur: there is significant staff turnover, staff not using the information in the way it was intended, or a there is a significant change to the PRM that staff need to understand.

Ensuring that PRMs are explainable and transparent: To promote public trust when implementing a new or modified PRM, it is also critical that agencies are transparent about how the PRM was developed and how it will be used. Part of being transparent requires child welfare agencies to provide some information explaining the model. As discussed further in the section on transparency and explainability, there is no agreed-upon standard for what information needs to be provided for a model to be explainable; however, at a basic level, explainability requires that the agency provide some information about what data were used to develop the PRM, the modeling approach used, and what the PRM was designed to predict.

Monitoring the Performance of a PRM

After a PRM is implemented, it is important for child welfare agencies to continue monitoring its performance, as well as its impact on child welfare decision making and outcomes for families and communities. When monitoring the performance, use, and impact of the PRM, agencies can take the following actions to reduce the risk it will exacerbate racial or ethnic bias in child welfare decision making:

- Continuing to examine the predictive performance of the PRM for racial and ethnic minority groups
- Evaluating its impact on equity indicators and other important outcomes
- Engaging community members, advocates, and end users in ongoing decisions about how to use the PRM

Monitoring predictive performance: The predictive performance of a PRM will change over time as child welfare practices change and populations or other features in the data sets shift (Drake et al., 2020; Matheny et al., 2019, p. 166). To ensure that a PRM continues to perform well across racial and ethnic subgroups, child welfare agencies should examine its predictive performance overall, as well as for specific subgroups, at defined time intervals and when there are recognized changes in child welfare policies or the data sets used to develop the PRM. Several techniques can be used to detect "data drift" or changes in the underlying feature distributions for the variables or features included in the model, such as sequential analysis methods that evaluate changes in error rates to determine whether data drift has occurred (Davies et al., 2020; Oladele, 2022). Agencies and developers should discuss the appropriate time intervals for monitoring the PRM's performance and techniques for identifying data drift and model performance.

If predictive performance degrades over time overall or for a specific subgroup, agencies may need to regenerate or retrain the PRM, re-examine the choice of the outcome variable and label on which the algorithm is trained, or recalibrate the PRM (Matheny et al., 2019; Obermeyer et al., 2019).

Evaluating the impact of a PRM: In addition to understanding if and how the predictive performance of a PRM changes over time once the tool has been implemented, it is also important to evaluate the tool's impact on child welfare decisions and important outcomes, such as the accuracy or consistency of screening referrals (Goldhaber-Fiebert & Prince, 2019; Lee et al., 2019). It is especially important for child welfare agencies to understand the impact that a PRM has on important indicators of equity identified in collaboration with community members, advocates, and regulators. For example, these agencies should understand whether the impact of the PRM differed for families from different racial and ethnic minority groups (Exhibit F) (Goldhaber-Fiebert & Prince, 2019). Agencies may benefit from partnering with external researchers to conduct rigorous evaluations unless they have the internal expertise and resources to conduct the evaluations themselves (Reisman et al., 2018).

Exhibit F. Evaluating the Allegheny Family Screening Tool (AFST) on equity and other outcomes

The Allegheny County Department of Human Services (DHS) implemented the AFST PRM to enhance its child welfare call screening decision-making process. In 2015, the Allegheny County DHS contracted with two separate research groups to evaluate the implementation and impact of the AFST (Allegheny County, 2022).

Hornby Zeller Associates, Inc. examined the implementation process based on interviews and surveys with staff and external stakeholders. These researchers found that end users appreciated DHS transparency regarding the tool's development and implementation, but that less than 50% of staff felt that the PRM improved the screening process (Hornby Zeller Associates, Inc., 2018).

Researchers from Stanford University used a pre/post design to examine the AFST's impact on key outcomes, including the overall rate of children screened in for investigation, the likelihood that children screened out would have no re-referrals within two months, and the likelihood that a child who was screened in had a case opened for services upon investigation or had a re-referral within two months. The researchers also looked at the differences in these outcomes for racial and ethnic minority groups. They found that using the AFST increased the rate of children screened in who had a case opened for services; had no impact on the re-referral rates among screened-out children, and led to reductions in disparities of rates of opening new cases between Black and White children (Goldhaber-Fiebert & Prince, 2019).

Reports from both evaluations were made publicly available, and the Allegheny County of DHS used the results to improve the AFST and its use by child welfare staff.

Agencies should also consider implementing the PRM in a way that facilitates a high-quality evaluation (Matheny et al., 2019, p. 165). This process could include, for example, a stepwise implementation, in which some staff begin using the PRM earlier than others, enabling an examination of differences in decision making between staff using the tool and those not yet doing so, or implementing the PRM as part of a randomized controlled trial (RCT), as was done in Douglas County, Colorado (Fitzpatrick & Wildeman, 2021).

Depending on the approach to implementation and resources available, possible methods for evaluating PRMs include step-wedge designs, RCTs, pre/post evaluations to identify changes to decision-making patterns that occurred after a PRM was implemented, or qualitative or survey evaluations designed to understand how agency staff are using a PRM and how it affects staff and those impacted by it (Goldhaber-Fiebert & Prince, 2019; Hornby Zeller Associates, Inc., 2018; Matheny et al., 2019). To promote transparency, child welfare agencies should publish the methods used to evaluate a PRM, the findings of that evaluation, and a description of any changes made to the PRM or its use based on the evaluation findings (Goldhaber-Fiebert & Prince, 2019; Hornby Zeller Associates, Inc., 2018). Ideally, agencies should also ensure that the public has the opportunity to comment on the findings and any changes made as a result.

Continuing to engage community members, advocates, and other end users in decisions about the continued use of PRMs: Child welfare agencies should continue to engage community members, advocates, child welfare professionals, and regulators during the monitoring stage as they consider concerns related to the ongoing use of PRMs (Elgin, 2018). Agencies should inform these groups about ongoing approaches for monitoring a PRM's performance and evaluating its impact. They could also be engaged in conversations about the appropriate design for the evaluation or selecting external researchers with whom to partner for evaluation activities. Finally, if issues are identified based on ongoing monitoring or the evaluation results, it is important to engage community members and others in decisions about how to address these issues, which may range from abandoning the use of the PRM to revising the model to updating policies related to how its predictions are used, depending on the nature of the issue identified. Roundtable participants suggested that, to promote transparency, child welfare agencies also might want to develop publicly facing documents addressing frequently asked questions that demonstrate the steps agencies took to address issues and concerns raised during the process of monitoring the performance and impact of PRMs.

ENSURING THAT PRMS ARE EXPLAINABLE AND THEIR USE IS TRANSPARENT

> *"If there is no transparent information on how algorithms and processes work, it is almost impossible to evaluate the fairness of the algorithms or discover discriminatory patterns in the system."*
>
> – Favaretto et al. (2019)

Ensuring that PRMs are explainable and their development and use are transparent is critical in promoting public oversight, accountability, and avenues for appeal (Favaretto et al., 2019; Pittsburgh Task Force on Public Algorithms, 2022; Samant et al., 2021; The Annie E. Casey Foundation, 2020; Yen & Hung, 2021). Explainable PRMs used transparently can also (1) encourage buy-in from professional staff, in turn improving the relationship decision makers have with the algorithm and resulting in better data quality and model performance; (2) promote continuous quality improvement; and (3) inform decisions about adopting or abandoning PRMs (Zytek et al., 2021).

Transparency requires that PRM developers and implementers publicly share information about how data on people are being used and how agencies are using these models to inform child welfare decisions (Dare, 2018; Samant et al., 2021; The Annie E. Casey Foundation, 2020). Engaging community members, advocates, and end users throughout the process of developing and deploying a PRM is an important step in promoting transparency. Additionally, child welfare agencies should be prepared to explain to affected families how they make certain decisions, including the types of information they use. A roundtable participant suggested that if a risk score from a PRM is used to determine whether an investigation should be conducted or an intervention implemented, the family should receive the risk score and information about the factors that influenced the score, and be given the opportunity to question the PRM's application to their particular case. However, if the risk score is just one piece of information used to inform the decision, explaining the decision-making approach is likely sufficient.

There is no agreed-upon standard regarding what is required for a PRM to be considered "explainable," but at a basic level, explainability requires that people using or affected by a PRM have some understanding of the model inputs, the methodology used to develop it, and its outputs (Lanier et al., 2020). To the extent feasible, child welfare agencies and developers should also describe the predictor variables or features that significantly influenced the prediction or outcome. Providing this information can enable the public to see how the PRM will interpret their behavior and choices, and, as a result, how agencies will act in future scenarios and what families can change to prevent future child welfare involvement (Favaretto et al., 2019). However, with complex algorithms, such as those using random forest methods or other advanced statistical techniques, it may not always be possible to explain why certain predictor variables are important and how they factor into the model output, which can create challenges for child welfare agencies attempting to explain why certain predictor variables influenced the model outcome (Chouldechova et al., 2018). Understanding what information about a PRM is most useful for people using and affected by it is an ongoing area of research and likely differs depending on the specific PRM use case.

Nevertheless, for all PRMs, to promote transparency and provide a basic level of information about the model itself, child welfare agencies should, at a minimum, publish a public report describing their methodology for creating a PRM that includes the data used; analytic documentation, including the identification of all predictor and outcome variables; how the model is being used; and the model performance details, including the results of evaluations and monitoring activities. The report should be publicly accessible and understandable to a variety of audiences, including agency staff, advocates, and community members (Chapin Hall and Chadwick

Center, 2018; Cuccaro-Alamin et al., 2017; Pittsburgh Task Force on Public Algorithms, 2022). As mentioned previously, it may also be useful to publish a frequently asked questions document that includes information about any steps the agency took to improve the PRM based on community feedback or the results from monitoring and evaluation activities. To ensure that the report and any other supporting materials are accessible, developers and implementers should engage community members, advocates, regulators, and child welfare professionals in determining how to share information about the PRM with people it might impact (Yen & Hung, 2021). Finally, as mentioned previously, child welfare agencies should work only with vendors who agree to waive proprietary information regarding the development and validation of a PRM (Cuccaro-Alamin et al., 2017; Drake et al., 2020; Francis et al., 2022; The Annie E. Casey Foundation, 2020).

STEPS FEDERAL AGENCIES CAN TAKE TO PROMOTE EQUITY IN PROJECTS USING PREDICTIVE ANALYTICS

Although not all the actions suggested in this report are relevant to every PRM developed to inform child welfare decisions, federal agencies with overlapping service populations and interests in PRMs can take a number of steps to help ensure that organizations receiving federal funds take appropriate actions to mitigate bias in any of their projects relying on predictive analytics. These agencies include the Administration for Children & Families, the Substance Abuse and Mental Health Services Administration, the Centers for Disease Control and Prevention, the Office of the Assistance Secretary for Planning and Evaluation, and the Office of Civil Rights, among others. The steps described below are not to be viewed sequentially, but rather as interdependent in informing equity-focused policy. In addition, they reflect core themes modeled throughout the report: accountability, transparency, and the need for government and public engagement.

a. **Avoid supporting the development or use of proprietary PRMs that are not transparent or explainable, as they inhibit public oversight and government accountability.** Transparency and some level of explainability are critical in promoting public trust in PRMs. Just as child welfare agencies should avoid working with vendors that refuse to waive proprietary or trade secrecy information, federal agencies might want to avoid providing financial support to organizations unwilling to share information about their model, the approach taken to develop it, and information related to its predictive performance. Instead, federal agencies could support those organizations committed to inviting public scrutiny, promoting citizen data protection, and using third-party reviews to evaluate their PRMs.

b. **Develop guidelines highlighting the importance of making public methodology reports accessible to broad audiences as well as explainability standards that describe what information child welfare professionals and those affected by PRMs need to understand about how these models inform decision-making processes.** Similar to the previous step, making methodology reports publicly available is important in promoting transparency. Federal agencies could consider creating guidelines about what information should be included in methodology reports and standards for public readability and accessibility based on methodology reports that have been publicized to date, such as those published by the Allegheny County Department of Human Services and the Douglas County Department of Human Services (Fitzpatrick & Wildeman, 2021). Federal agencies could also consider the minimum requirements for ensuring that PRMs are at least partially explainable to end users and those affected by the model. Once these guidelines and standards are developed, federal agencies could require that organizations receiving federal funding for projects using predictive analytics adhere to those guidelines and standards as a condition of their funding.

c. **Promote the engagement of community members, advocates, regulators, and child welfare professionals in developing, implementing, and monitoring PRMs in child welfare settings.** This report highlights various decision points during which organizations developing PRMs can engage community members and end users. We also highlight some approaches that organizations can adopt to promote effective engagement, such as using case review exercises to incorporate community member and agency staff feedback during the process of developing a PRM. Although organizations are still developing best practices for effectively engaging community members and other end users, some level of engagement should be present in all projects that involve predictive analytics in supporting policy decisions. Federal agencies could consider requiring all organizations receiving federal funding for such projects to develop a plan for community member and end-user engagement and adhere to that plan throughout their projects. Federal agencies might also want to consider earmarking some of their funding to support engagement activities.

d. **Support partnerships between agencies and researchers to explore effective ways of governing the use of PRMs in collaboration with affected communities and advocates.** Because there are no current federal regulations governing the use of PRMs in child welfare settings, it is important for child welfare agencies and researchers or vendors to agree on approaches for effectively governing these models as they are developed and implemented. Governance models could include standards for the following: protecting data used to develop and validate PRMs, data quality, predictive performance and performance across subgroups, and monitoring PRMs once they are implemented. Once governance models have been developed, federal agencies could share them publicly for comment. The agencies could support these activities by, for example, providing funding to an organization to convene other relevant organizations to develop a shared governance structure that could be broadly implemented, or by requiring that organizations receiving federal funding share their governance plans and approach for establishing those plans.

e. **Work with the child welfare community to understand best practices for evaluating PRMs to ensure that they do not increase—and ideally, reduce—disparities in the communities in which they are implemented.** As discussed in this report, a variety of methodological approaches can be adopted to evaluate the impact of PRMs on important equity-related outcomes. Federal agencies could work with the child welfare community to understand which evaluation designs are most feasible to implement in different situations and what resources are required to successfully complete evaluation activities. Developing guidelines related to PRM evaluation could help child welfare agencies plan for evaluation activities when implementing a PRM and help them understand with whom they may need to partner to conduct a successful evaluation. The guidelines should also include information about how to report the results from an evaluation so they are accessible to a public audience.

f. **Support the use of PRMs that have an explicit goal of reducing disparities and informing, rather than replacing, child welfare agency staff (Cheng et al., 2022).** When considering whether to support a predictive analytics project, federal agencies could consider prioritizing funding for PRMs that have an explicit goal of reducing disparities, such as PRMs designed to improve the allocation of scarce resources. In addition, federal agencies might want to meet with the model developers to understand their vision regarding how their model will be used, and might want to ask about the developer's perception of the risks and benefits of the proposed approach. PRMs, when implemented with the goal of reducing bias, should generally be used to supplement and not replace human decision making. Agency staff should feel empowered to apply their expertise in specific cases and be given clear instructions on PRM capabilities. Federal agencies might want to be cautious about any model that will be used to make rather than just inform a decision. These agencies might also want to conduct their own assessment of the risks and potential benefits of the proposed approach before deciding whether to move forward with supporting the proposed project. Finally, agencies might want to ask researchers to provide detailed plans for training end users on the appropriate use of a model to ensure it will be used in the way intended.

CONCLUSION

PRMs have the potential to provide child welfare agencies with information that can improve the process of making decisions and allocating resources to support children, families, and communities. Although child welfare agency staff have considerable expertise and training in making complex and consequential decisions related to child well-being, they do not have the capacity on their own to sift through large data sets to understand patterns between contextual factors and outcomes they hope to avoid. PRMs can identify these relationships and provide information about how different contextual factors may influence overall risk. However, these models also have the potential to exacerbate racial and ethnic disparities if they are not planned for, developed, implemented, and monitored using approaches aimed at identifying and mitigating potential sources of bias at each step. They also have the potential to increase mistrust in CPS if child welfare agencies are not transparent about how they use or plan to use such models to inform the choices they make.

Throughout this report, we highlight actions that child welfare agencies can take over the course of integrating a PRM to reduce the risk that these models may exacerbate racial and ethnic disparities. The actions described in this report highlight the need to (1) carefully define the problem to be solved by a PRM and the specific PRM use case; (2) interrogate any data sets that will be used to develop a PRM to understand data quality issues and the potential for the data to reflect historical biases; (3) evaluate the PRM outcome, as well as the model's overall impact and performance; (4) assess the pros and cons of implementing a PRM and how best to balance misclassification rates; (5) train staff to use the PRM appropriately; and (6) monitor a PRM's performance following its implementation to determine whether its performance degrades over time. Throughout the process of integrating a PRM, it is also critical to engage community members, advocates, agency staff, and others to ensure the PRM meets their expectations, to build trust, and to ensure transparency. Taken together, these actions, especially when supported by government agencies funding the development of these models, can help ensure that PRMs used by child welfare agencies are fair, high performing, and achieve the outcomes they were designed to achieve.

REFERENCES

1. Allegheny County. (2022). *The Allegheny Family Screening Tool*. Allegheny County. https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx
2. Allen, A., Mataraso, S., Siefkas, A., Burdick, H., Braden, G., Dellinger, R.P., McCoy, A., Pellegrini, E., Hoffman, J., Green-Saxena, A., Barnes, G., Calvert, J., & Das, R. (2020). A racially unbiased, machine learning approach to prediction of mortality: Algorithm development study. *JMIR Public Health and Surveillance*, *6*(4), e22400. https://doi.org/10.2196/22400
3. Baird, C., Wagner, D., Healy, T., & Johnson, K. (1999). Risk assessment in child protective services: Consensus and actuarial model reliability. *Child Welfare*, *78*(6), 723–748.
4. Barocas, S., & Boyd, D. (2017). Engaging the ethics of data science in practice. *Communications of the ACM*, *60*(11), 23–25. https://doi.org/10.1145/3144172
5. Cahan, E.M., Hernandez-Boussard, T., Thadaney-Israni, S., & Rubin, D.L. (2019). Putting the data before the algorithm in big data addressing personalized healthcare. *NPJ Digital Medicine*, *2*, 78. https://doi.org/10.1038/s41746-019-0157-2
6. Capatosto, K. (2017, May 31). Accounting for racial bias when applying predictive analytics to child welfare. *The Imprint*. https://imprintnews.org/opinion/accounting-racial-bias-applying-predictive-analytics-child-welfare/27074
7. Casey Family Programs. (2018). How does investigation, removal, and placement cause trauma for children? Casey.org. https://caseyfamilypro-wpengine.netdna-ssl.com/media/SC_Investigation-removal-placement-causes-trauma.pdf
8. Chapin Hall and Chadwick Center. (2018). Making the most of predictive analytics: Responsible and innovative uses in child welfare policy and practice. Children's San Diego and the University of Chicago. https://www.chapinhall.org/wp-content/uploads/Making-the-Most-of-Predictive-Analytics.pdf
9. Cheng, H.-F., Stapleton, L., Kawakami, A., Sivaraman, V., Cheng, Y., Qing, D., Perer, A., Holstein, K., Wu, Z.S., & Zhu, H. (2022). How child welfare workers reduce racial disparities in algorithmic decisions. *CHI Conference on Human Factors in Computing Systems*, 1–22. https://doi.org/10.1145/3491102.3501831
10. Child Welfare Information Gateway. (2021). Child welfare practice to address racial disproportionality and disparity (Bulletins for Professionals). U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau. https://www.childwelfare.gov/pubs/issue-briefs/racial-disproportionality/
11. Chouldechova, A., & G'Sell, M. (2017). Fairer and more accurate, but for whom? *ArXiv Preprint*. https://doi.org/10.48550/arXiv.1707.00046
12. Chouldechova, A., Putnam-Hornstein, E., Benavides-Prado, D., Fialko, O., & Vaithianathan, R. (2018). A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. *Proceedings of Machine Learning Research*, 134–148. https://www.cs.ubc.ca/~conati/522/532b-2019/papers/chouldechovaCaseStudyPredictionFairnessEtc.pdf
13. Cuccaro-Alamin, S., Foust, R., Vaithianathan, R., & Putnam-Hornstein, E. (2017). Risk assessment and decision making in child protective services: Predictive risk modeling in context. *Children and Youth Services Review*, *79*, 291–298.
14. Damman, J.L., Johnson-Motoyama, K., Wells, S., & Harrington, K. (2020). Factors associated with the decision to investigate child protective services referrals: A systematic review. *Child & Family Social Work, 25*, 785 – 804.
15. Dare, T. (2018). *Ethical evaluation of the Predict-Align-Prevent Program*. The University of Auckland. https://www.predict-align-prevent.org/_files/ugd/fbb580_0bf866ed131c4e8b8614c809995c6f06.pdf
16. Davies, N.G., Klepac, P., Liu, Y., Prem, K., Jit, M., & Eggo, R.M. (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine*, *26*(8), 1205–1211. https://doi.org/10.1038/s41591-020-0962-9

17. Drake, B., & Jonson-Reid, M. (2018). Administrative data and predictive risk modeling in public child welfare: Ethical issues relating to California. Brown School of Social Work, Washington University in St. Louis. https://www.datanetwork.org/wp-content/uploads/ethical-review-of-predictive-risk-modeling-1.pdf
18. Drake, B., Jonson-Reid, M., Ocampo, M.G., Morrison, M., & Dvalishvili, D. (2020). A practical framework for considering the use of predictive risk modeling in child welfare. *The Annals of the American Academy of Political and Social Science*, *692*(1), 162–181.
19. Elgin, D.J. (2018). Utilizing predictive modeling to enhance policy and practice through improved identification of at-risk clients: Predicting permanency for foster children. *Children and Youth Services Review*, *91*, 156–167. https://doi.org/10.1016/j.childyouth.2018.05.030
20. Favaretto, M., De Clercq, E., & Elger, B.S. (2019). Big data and discrimination: Perils, promises and solutions. A systematic review. *Journal of Big Data*, *6*(1), 12. https://doi.org/10.1186/s40537-019-0177-4
21. Feng, A., & Wu, S. (2019, May 1). The myth of the impartial machine. *Parametric Press*, *1*. https://parametric.press/issue-01/the-myth-of-the-impartial-machine/
22. Fitzpatrick, M.D., & Wildeman, C. (2021). Final report on Douglas County Decision Aid (DCDA) predictive risk modeling randomized control trial experiment. https://csda.aut.ac.nz/__data/assets/pdf_file/0012/504102/Douglas-RCT-Final-Report-210211.pdf
23. Francis, C., Froomkin, D., Pales, E., Sung, K., & Wooten, K. (2022). Algorithmic accountability: The need for a new approach to transparency and accountability when government functions are performed by algorithms. Media Freedom & Information Access Clinic, Yale Law School. https://law.yale.edu/sites/default/files/area/center/mfia/document/algorithmic_accountability_report.pdf
24. Goldhaber-Fiebert, J.D., & Prince, L. (2019). Impact evaluation of a predictive risk modeling tool for Allegheny County's Child Welfare Office [Evaluation Report]. https://www.alleghenycountyanalytics.us/wp-content/uploads/2019/05/Impact-Evaluation-from-16-ACDHS-26_PredictiveRisk_Package_050119_FINAL-6.pdf
25. Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29–36. https://doi.org/10.1148/radiology.143.1.7063747
26. Hornby Zeller Associates, Inc. (2018). Allegheny County predictive risk modeling tool implementation: Process evaluation. https://www.alleghenycounty.us/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=6442467253
27. Kennedy, E.E., Bowles, K.H., & Aryal, S. (2021). Systematic review of prediction models for postacute care destination decision-making. *Journal of the American Medical Informatics Association*, *29*(1), 176–186. https://doi.org/10.1093/jamia/ocab197
28. Kim, H., Wildeman, C., Jonson-Reid, M., & Drake, B. (2017). Lifetime prevalence of investigating child maltreatment among US children. *American Journal of Public Health*, *107*(2), 274–280. https://doi.org/10.2105/AJPH.2016.303545
29. Krakouer, J., Wu Tan, W., & Parolini, A. (2021). Who is analysing what? The opportunities, risks and implications of using predictive risk modelling with Indigenous Australians in child protection: A scoping review. *Australian Journal of Social Issues (John Wiley & Sons, Inc.)*, *56*(2), 173–197. aph. https://doi.org/10.1002/ajs4.155
30. Lanier, P., Rodriguez, M., Verbiest, S., Bryant, K., Guan, T., & Zolotor, A. (2020). Preventing infant maltreatment with predictive analytics: Applying ethical principles to evidence-based child welfare policy. *Journal of Family Violence*, *35*(1), 1–13.
31. Lee, N.T., Resnick, P., & Barton, G. (2019). *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*. Brookings Institution. https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

32. Matheny, M., Thadaney Israni, S., & Ahmed, M. (2019). *Artificial intelligence in health care: The hope, the hype, the promise, the peril* (The Learning Health Systems Series). NAM Special Publication. https://nam.edu/artificial-intelligence-special-publication/
33. Metz, A. (2016). Practice profiles: A process for capturing evidence and operationalizing innovations [White Paper]. National Implementation Research Network, University of North Carolina at Chapel Hill. https://nirn.fpg.unc.edu/resources/white-paper-practice-profiles-process-capturing-evidence-and-operationalizing-innovations
34. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342
35. Oladele, S. (2022, July 21). A comprehensive guide on how to monitor your models in production. *Neptune Labs*. https://neptune.ai/blog/how-to-monitor-your-models-in-production-guide
36. Oregon DHS. (2019). *Safety at screening tool development and execution report*. Oregon Department of Human Services, Office of Reporting, Research, Analytics and Implementation (ORRAI). https://www.oregon.gov/dhs/ORRAI/Documents/safety-at-screening-report.pdf
37. Park, Y., Hu, J., Singh, M., Sylla, I., Dankwa-Mullan, I., Koski, E., & Das, A.K (2021). Comparison of methods to reduce bias from clinical prediction models of postpartum depression. *JAMA Network Open*, *4*(4), e213909–e213909. c8h. https://doi.org/10.1001/jamanetworkopen.2021.3909
38. Pittsburgh Task Force on Public Algorithms. (2022). *Report of the Pittsburgh Task Force on Public Algorithms*. Institute for Cyber Law, Policy, and Security, University of Pittsburgh. https://www.cyber.pitt.edu/sites/default/files/pittsburgh_task_force_on_public_algorithms_report.pdf
39. Pryce, J., Yelick, A., Zhang, Y., & Fields, K. (2018). Using artificial intelligence, machine learning, and predictive analytics in decision-making. Florida Institute for Child Welfare, Florida State University. https://ficw.fsu.edu/sites/g/files/upcbnu1106/files/Final%20Reports/FICW%20Using%20Artificial%20Intelligence,%20Machine%20Learning,%20and%20Predictive%20Analytics%20in%20Decision-Making.pdf
40. Purdy, J., & Glass, B. (2020). The pursuit of algorithmic fairness: On “correcting” algorithmic unfairness in a child welfare reunification success classifier. https://doi.org/10.48550/arXiv.2010.12089
41. Rajkomar, A., Hardt, M., Howell, M.D., Corrado, G., & Chin, M.H. (2018). Ensuring fairness in machine learning to advance health equity. *Annals of Internal Medicine*, *169*(12), 866–872. https://doi.org/10.7326/M18-1990
42. Reisman, D., Schultz, J., Crawford, K., & Whittaker, M. (2018). Algorithmic impact assessments: A practical framework for public agency accountability (AI Now Institute Report). New York University. https://ainowinstitute.org/aiareport2018.pdf
43. Samant, A., Horowitz, A., Xu, K., & Beiers, S. (2021). *Family surveillance by algorithm*. American Civil Liberties Union. https://www.aclu.org/fact-sheet/family-surveillance-algorithm
44. Shlonsky, A., & Wagner, D. (2005). The next step: Integrating actuarial risk assessment and clinical judgment into an evidence-based practice framework in CPS case management. *Children and Youth Services Review*, 27(4), 409–427. https://doi.org/10.1016/j.childyouth.2004.11.007
45. The Annie E. Casey Foundation. (2020). Four principles to make advanced data analytics work for children and families. The Annie E. Casey Foundation. https://assets.aecf.org/m/resourcedoc/aecf-fourprinciplestomakeadvanced-2020.pdf
46. U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children’s Bureau. (2021). *Child Maltreatment 2019*.
47. U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children’s Bureau. (2022a). *Child Maltreatment 2020.* https://www.acf.hhs.gov/sites/default/files/documents/cb/cm2020.pdf
48. U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children’s Bureau. (2022b). *Child Welfare Outcomes 2019: Report to Congress. Executive Summary*. https://www.acf.hhs.gov/sites/default/files/documents/cb/cwo-executive-summary-2019.pdf

49. Vaithianathan, R., Dinh, H., Kalisher, A., Kithulgoda, C., Kulick, E., Mayur, M., Ning, A., Benavides Prado, D., & Putnam-Hornstein, E. (2019). *Implementing a child welfare decision aide in Douglas County* [Methodology Report]. Centre for Social Data Analytics. https://csda.aut.ac.nz/__data/assets/pdf_file/0009/347715/Douglas-County-Methodology_Final_3_02_2020.pdf
50. Vaithianathan, R., Kulick, E., Putnam-Hornstein, E., & Benavides Prado, D. (2019). *Allegheny Family Screening Tool: Methodology, Version 2*. https://www.alleghenycountyanalytics.us/wp-content/uploads/2019/05/Methodology-V2-from-16-ACDHS-26_PredictiveRisk_Package_050119_FINAL-7.pdf
51. Vaithianathan, R., Putnam-Hornstein, E., Stagner, M., & Weigensburg, E. (2019). *Predictive risk modeling for child protection* [Child Welfare Fact Sheet]. Mathematica. https://www.mathematica.org/publications/predictive-risk-modeling-for-child-protection
52. Yen, C.P., & Hung, T.W. (2021). Achieving equity with predictive policing algorithms: A social safety net perspective. *Science and Engineering Ethics*, *27*(3), 36. https://doi.org/10.1007/s11948-021-00312-x
53. Zytek, A., Liu, D., Vaithianathan, R., & Veeramachaneni, K. (2021). Sibyl: Explaining machine learning models for high-stakes decision making. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3411763.3451743

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Office of the Assistant Secretary for Planning and Evaluation

200 Independence Avenue SW, Mailstop 447D
Washington, D.C. 20201

For more ASPE briefs and other publications, visit:
aspe.hhs.gov/reports

ABOUT THE AUTHORS
Danielle Whicher and Allon Kalisher are senior researchers at Mathematica; *Emma Pendl-Robinson* is a data scientist at Mathematica; and *Kyla Jones* is a project associate at Mathematica.

SUGGESTED CITATION
Whicher, D., Pendl-Robinson, E., Jones, K. and Kalisher, A. Avoiding Racial Bias in Child Welfare Agencies' Use of Predictive Risk Modeling. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. October 2022.

DISCLOSURE
This communication was printed, published, or produced and disseminated at U.S. taxpayer expense.

Subscribe to ASPE mailing list to receive
email updates on new publications:
https://list.nih.gov/cgi-bin/wa.exe?SUBED1=ASPE-HEALTH-POLICY&A=1

For general questions or general
information about ASPE:
aspe.hhs.gov/about