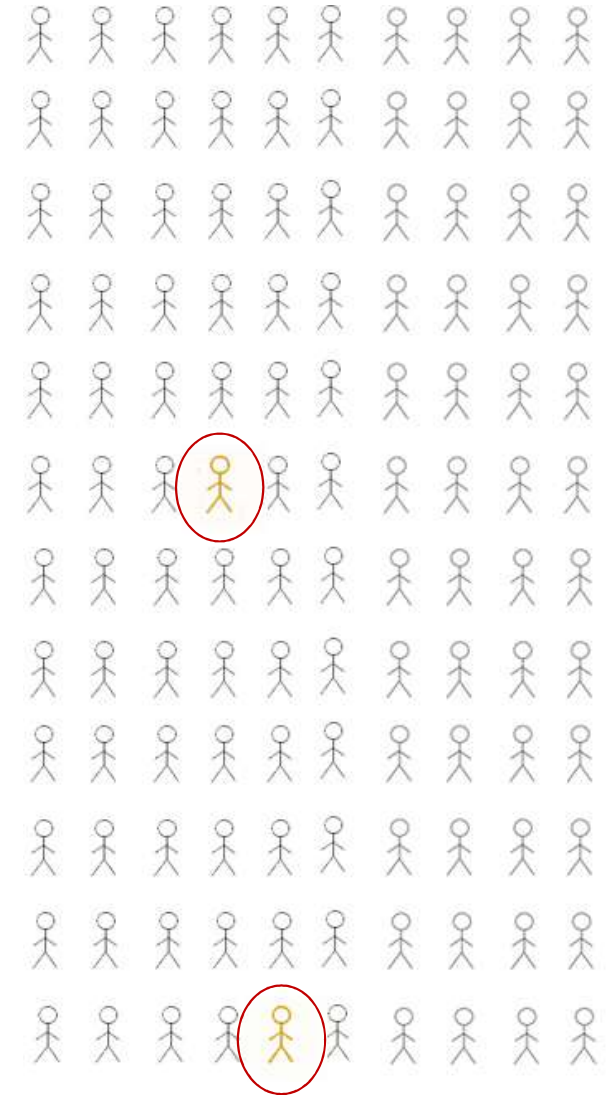# Prediction as a Tool for Prevention: Developing and Applying Algorithms in Child Protection Systems

June 2019

overview

- Predictive Risk Modeling
- Child Protection Context
- Example...California
  - * Model build
  - * Value Add
  - * Equity
  - * Validation

Predictive Risk Modeling

predictive risk modeling

- Automated tool / model / process by which a score is generated from existing information (*predictors*) and reflects the likelihood that some future event *(outcome)* will occur.

- Risk stratification allows for greater efficiency and effectiveness (*i.e., where can we best direct resources to prevent the adverse event*)

high risk

low risk
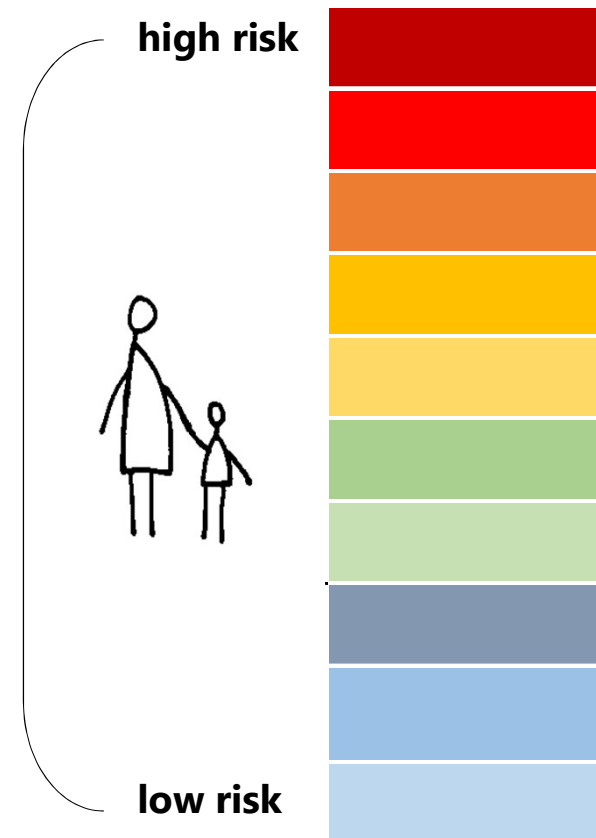
predictive risk modeling

Wider availability of data + advances in technologies + appreciation that humans are often poor at weighing a multitude of factors simultaneously...

opportunities

- Cost effective screening of large populations
- No subjective / human element
- Continuum of scores
- Models can be built using data from local populations
- Opportunities to reduce costs / improve performance by identifying high service utilizers

challenges

- Cost effective screening of large populations
- No subjective / human element
- Continuum of scores
- Models can be built using data from local populations
- Opportunities to reduce costs / improve performance by identifying high service utilizers

Child Welfare Context

lifetime

Los Angeles Times — Science Now
Discoveries from the world of science and medicine

1 in 8 U.S. children will become victim of serious abuse or neglect

---

☰ | **Home** | **Articles** | **Authors** | **Subscrip**

Lifetime Prevalence of Investigating Child Maltreatment Among US Children

Hyunil Kim MSW, **Christopher Wildeman** PhD, **Melissa Jonson-Reid** PhD, MSW, and **Brett Drake** PhD, MSW

[+] Author affiliations, information, and correspondence details

Abstract   Full Text   References   Supplements   PDF   PDF Plus

Objectives. To estimate the lifetime prevalence of official investigations for child maltreatment among children in the United States.

Methods. We used the National Child Abuse and Neglect Data System Child Files (2003–2014) and Census data to develop synthetic cohort life tables to estimate the cumulative prevalence of reported childhood maltreatment. We extend previous work, which explored only confirmed rates of maltreatment, and we add new estimations of maltreatment by subtype, age, and ethnicity.

Allegheny Family Screening Tool

The New York Times Magazine

Home

Can an Algorithm Tell When Kids Are in Danger?

Child protective agencies are haunted when they fail to save kids. Pittsburgh officials believe a new data analysis program is helping them make better judgment calls.

By DAN HURLEY JAN. 2, 2018

Screening Score   Historical Screening Scores

Family Screening Score

Recalculate Screening Score

The purpose of the **Family Screening Score** is to use information collected by DHS and other partners to inform screening decisions. The **Family Screening Score** is calculated by integrating and analyzing hundreds of data elements on each person related to the referral to generate an overall **Family Screening Score**. The score predicts the long-term likelihood of re-referral, if the referral is screened out without an investigation, or home removal, if the referral is screened in for investigation.

If the **Family Screening Score** meets the threshold for "mandatory screen-in," the call must be investigated. In all other circumstances, the **Family Screening Score** provides additional information to assist the Call Screening Unit in making a call screening decision and should not replace clinical judgement.

The **Family Screening Score** is only intended to inform call screening decisions and is not intended to be used in making investigative or other child welfare decisions.
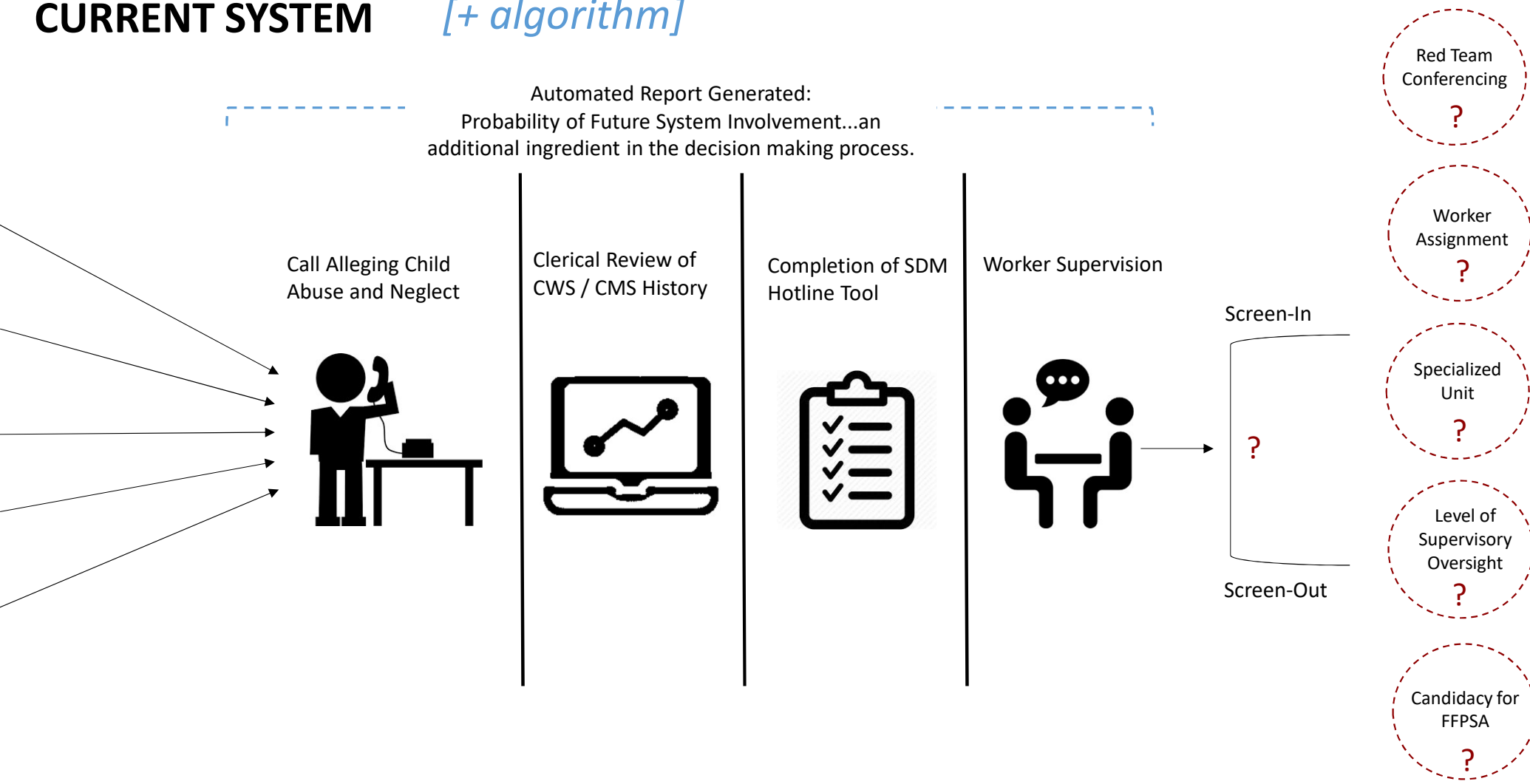
High

16

Moderate

Low

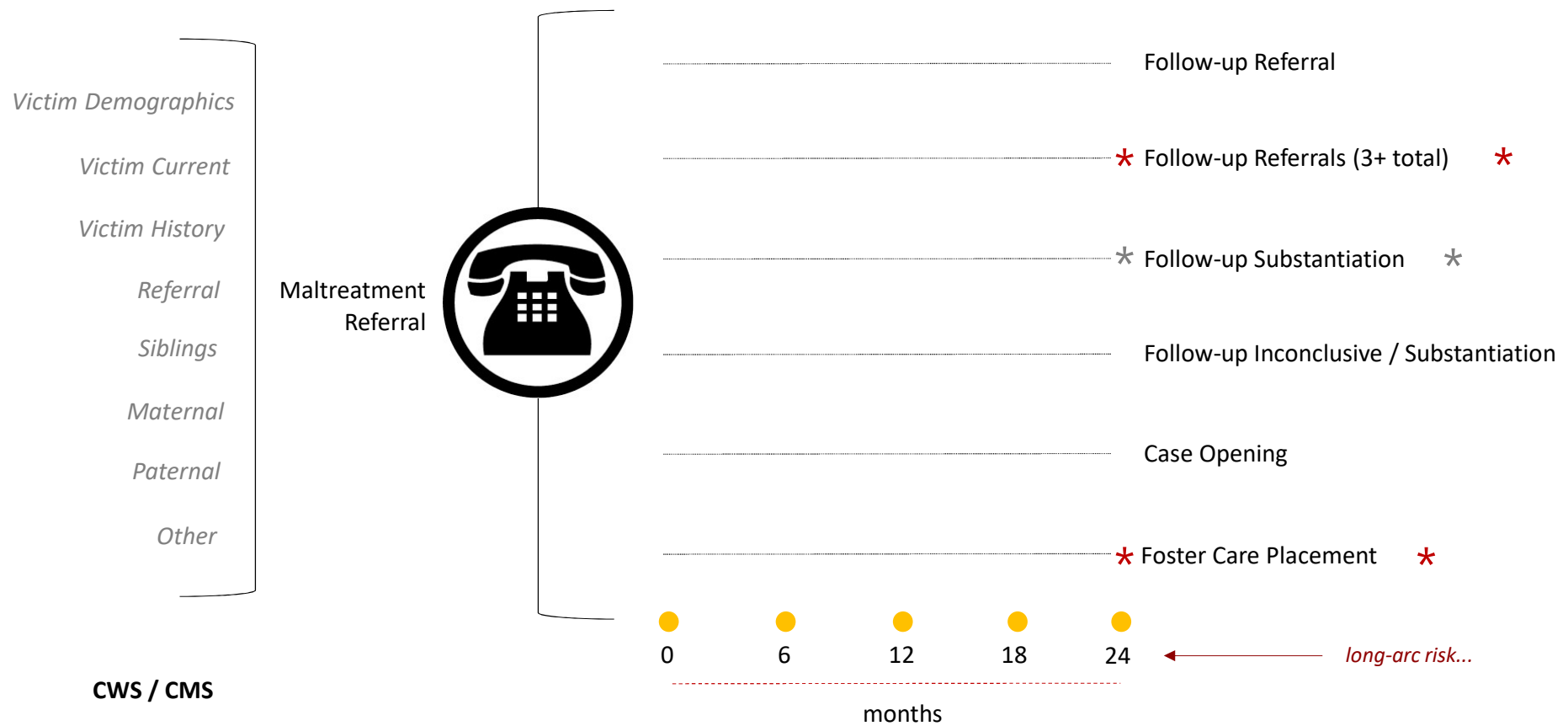**Last Run By:** Jane McBeth

Last Run Date: 4/7/2016, 10:32 AM

Algorithm Versions Used Re-referral v43 Placement v22
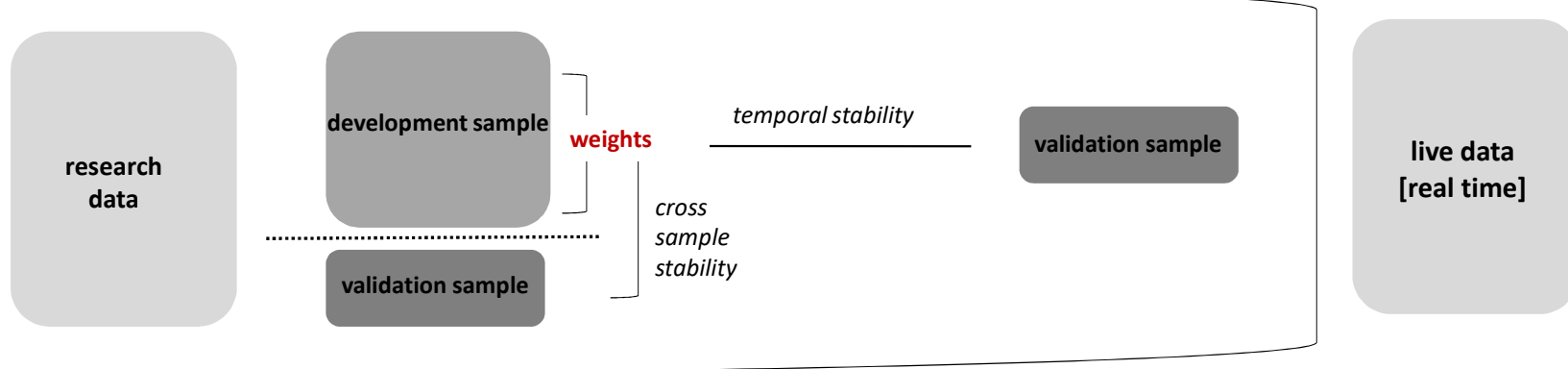
8

CURRENT SYSTEM _[+ algorithm]_

Automated Report Generated:
Probability of Future System Involvement...an
additional ingredient in the decision making process.

Call Alleging Child
Abuse and Neglect

Clerical Review of
CWS / CMS History

Completion of SDM
Hotline Tool

Worker Supervision

Screen-In

?

Screen-Out

Red Team
Conferencing
?

Worker
Assignment
?

Specialized
Unit
?

Level of
Supervisory
Oversight
?

Candidacy for
FFPSA
?

Scalability? Generalizability?

[system] OUTCOMES

Victim Demographics

Victim Current

Victim History

Referral

Siblings

Maternal

Paternal

Other

CWS / CMS

Maltreatment
Referral

.... Follow-up Referral

.... * Follow-up Referrals (3+ total)   *

.... * Follow-up Substantiation   *

.... Follow-up Inconclusive / Substantiation

.... Case Opening

.... * Foster Care Placement   *

| 0 | 6 | 12 | 18 | 24 |

← *long-arc risk...*

months

ALGORITHM DEVELOPMENT & VALIDATION

Features of the algorithm may not be easily interpreted by most people

But we can have transparency as to…

- *What goes into the model and where those data come from (predictors)*
- *What comes out of the model and how it performs across groups (accuracy)*
- *What the model is being used for (business use case)*
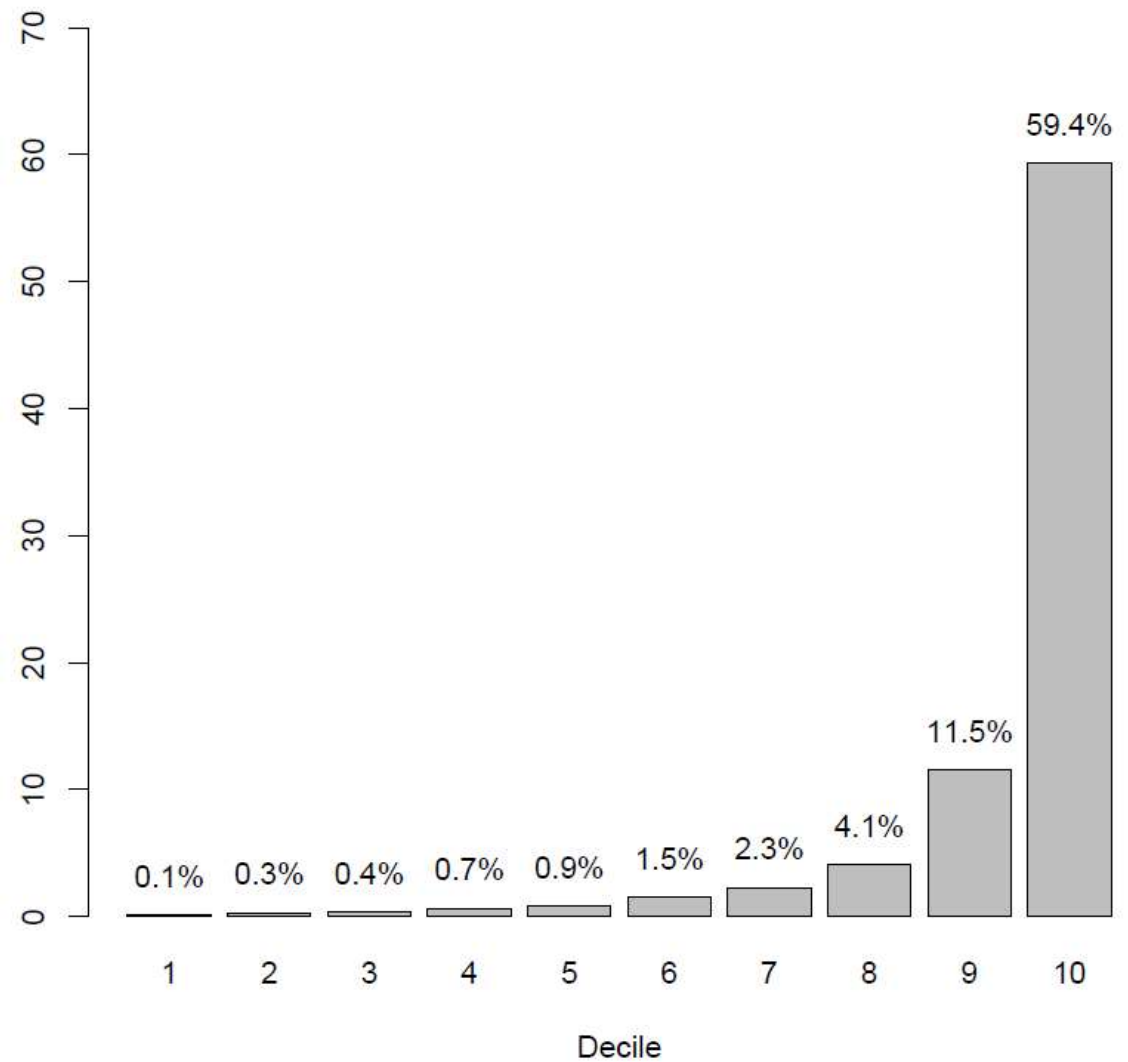
"risk deciles"

We use the algorithm to assign each child/referral into 10 evenly sized groups (or deciles) based on the predicted probability of system involvement.

Probability of System Involvement

Low ←——————————————————————————→ High

Decile	1	2	3	4	5	6	7	8	9	10
	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
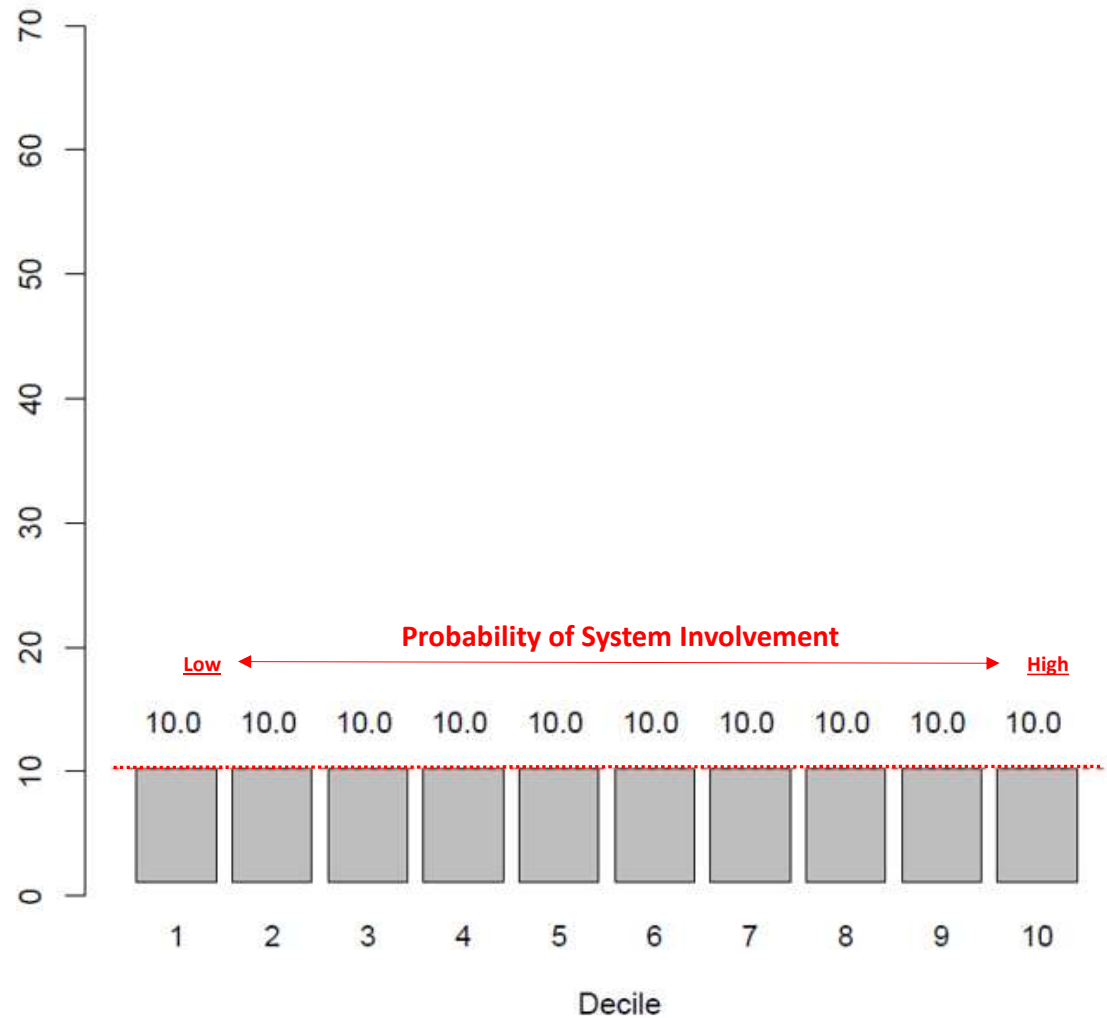
risk of foster care placement

We then examined how well the algorithm risk-stratified children (in unique family and referral events) by looking at how many children were placed in foster care within 24 months

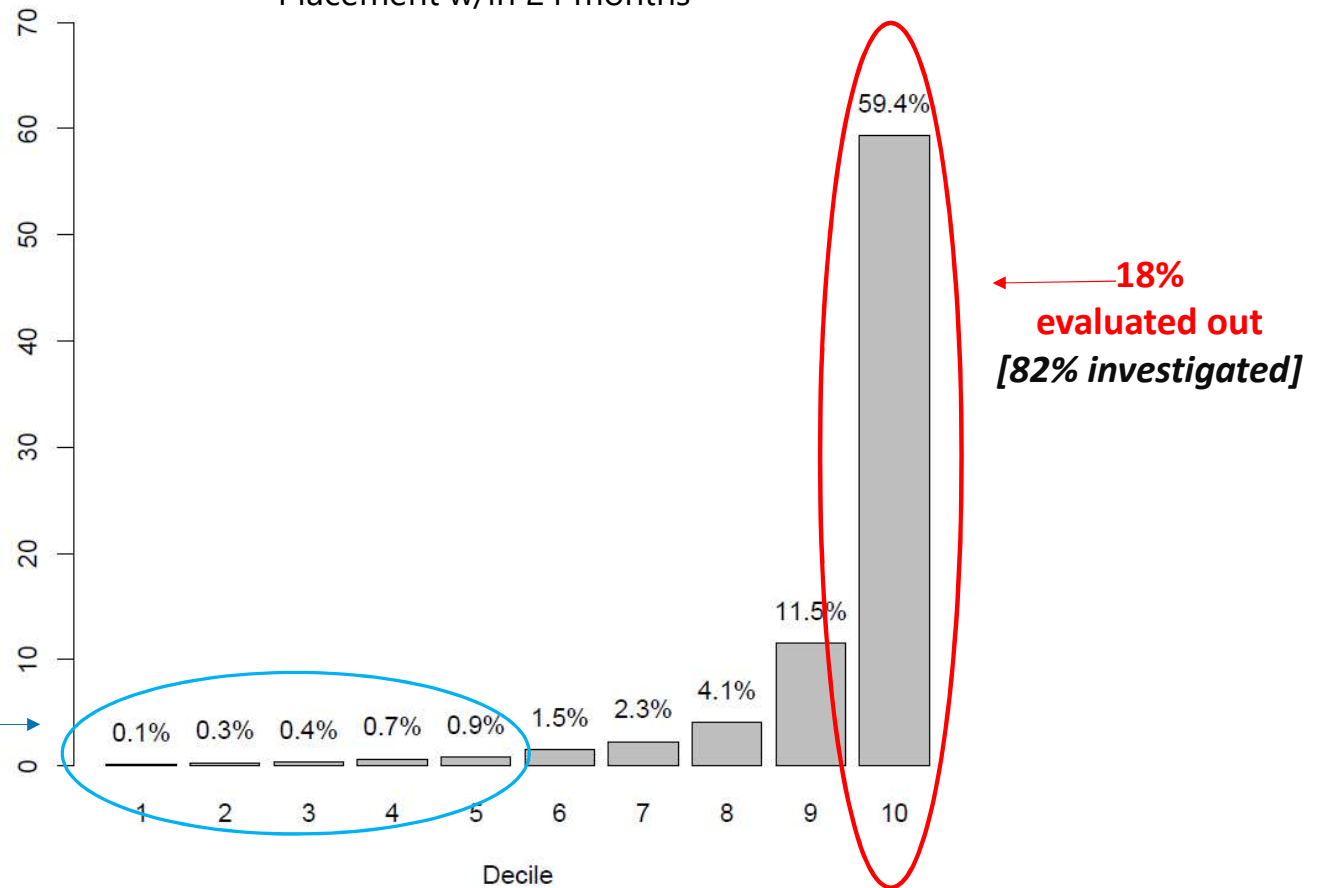(test set, statewide average: 8.0%)

Value Add..

"essentially, all models are wrong, but some are useful" (George E.P. Box, 1987)

historical screening decisions by risk decile

Test set only. Excludes children with a current open case or foster care placement.
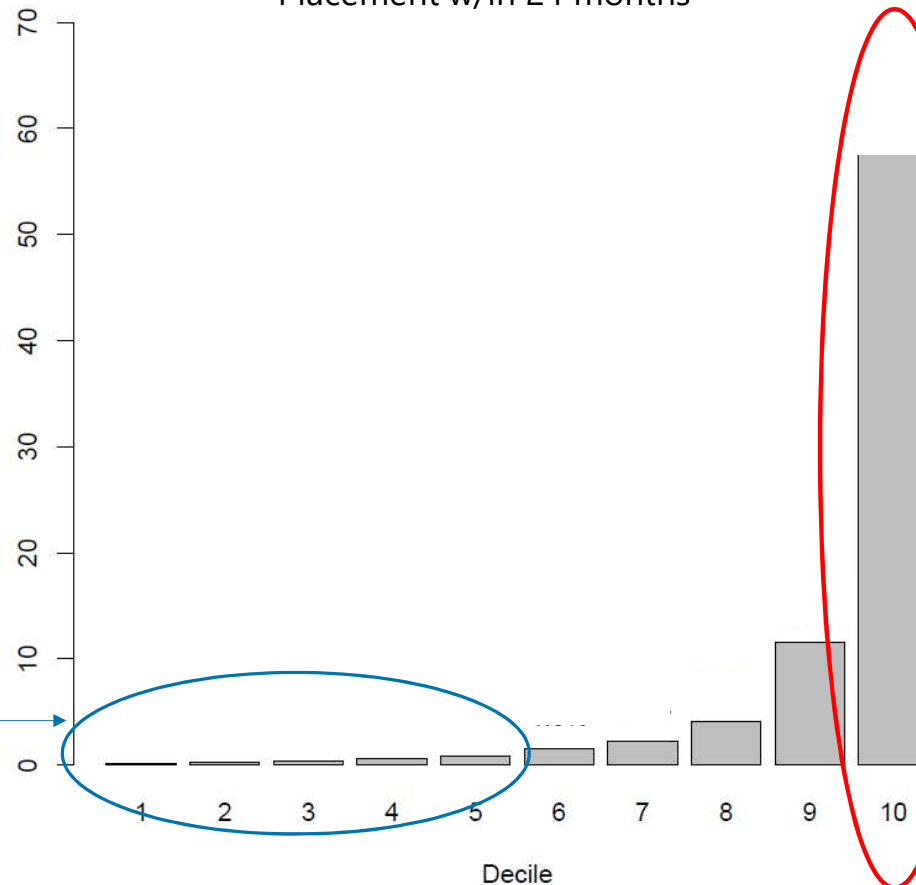
Placement w/in 24 months

37%
evaluated out
[63% investigated]

18%
evaluated out
[82% investigated]

placement outcomes?

Placement w/in 24 months

High Risk Screen-Outs?

18% of children in highest risk decile were screened out...

51% still ended up placed in foster care within 24 months

Low Risk Investigations?

63% of children in the lowest five risk deciles were investigated

Less than 1% were placed in foster care within 24 months

Decile

SDM® family risk assessment

SDM

"The SDM <u>family risk assessment</u> identifies families with low, moderate, high, or very high <u>probabilities of future abuse or neglect</u>. By completing the risk assessment, the worker obtains an <u>objective appraisal of the likelihood that a family will maltreat their child in the next 18 to 24 months</u>. The difference between risk levels is substantial. Families classified as high risk have significantly higher rates of subsequent referral and substantiation than families classified as low risk, and they are more often involved in serious abuse or neglect incidents."

http://www.childsworld.ca.gov/res/pdf/SDM_Manual.pdf
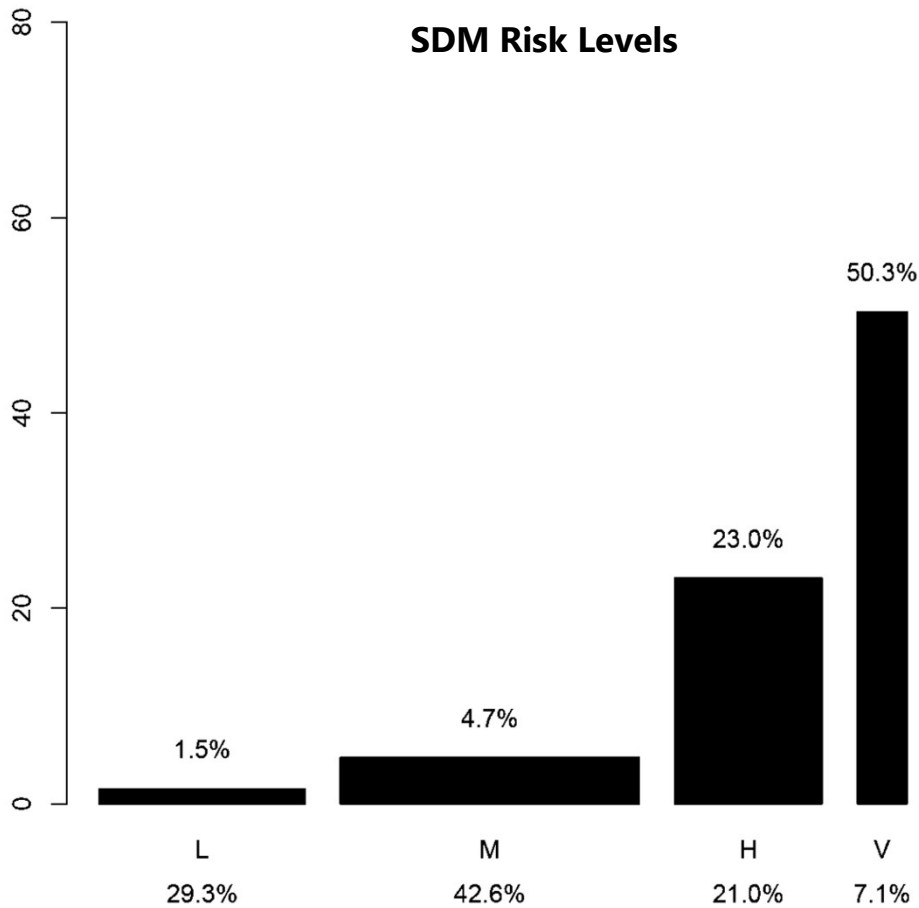
value add of an algorithm?

Best attempt at an "apples to apples" comparison...how would an algorithm improve supervision and practice above and beyond what we have today?

- We used risk scores generated by the algorithm to distribute children/families into groups equivalent in size to those observed through frontline staff's use of the SDM® family risk assessment
  - ✓ *Low Risk: 29.3%*
  - ✓ *Moderate Risk: 42.6%*
  - ✓ *High Risk: 21.0%*
  - ✓ *Very High Risk: 7.1%*

- We then examined classification decisions *(test sample, restricted to unique universe of 230,104 children/families)* for:
  - Outcome the algorithm was trained to predict *(i.e., foster care placement w/in 24m)*
  - Outcome the algorithm was trained to predict *(i.e., 3+ referrals w/in 24m)*
  - Outcome the SDM® risk assessment targets *(i.e., substantiated maltreatment w/in 24m)*

- We additionally examined algorithm classifications separately for children/families in the test set who had a completed SDM® risk assessment *(n=107,850)* vs. those that did not *(n=122,254)*, confirming that all findings held

- Overall, these comparisons examine practice conclusions that were drawn about future risk (defined by the completion and use of the SDM® risk assessment) versus those that would have been suggested by an algorithm
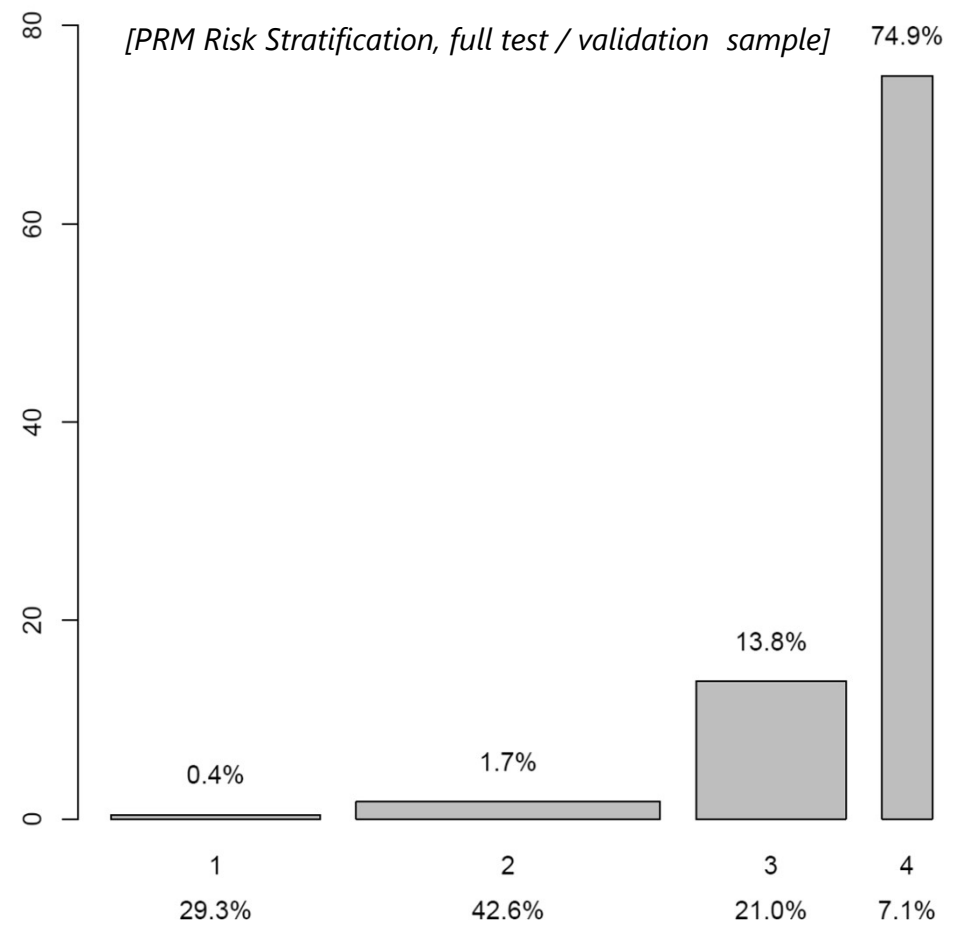
Risk stratification based on SDM risk assessment levels vs. PRM stratifications: Placements in Foster Care within 24 months

SDM Risk Levels

| L	M	H	V
1.5%	4.7%	23.0%	50.3%
29.3%	42.6%	21.0%	7.1%

[PRM Risk Stratification, full test / validation sample]

1	2	3	4
0.4%	1.7%	13.8%	74.9%
29.3%	42.6%	21.0%	7.1%

placement

Current use of the SDM risk assessment results in nearly half (42.6%) of children classified as "moderate risk".
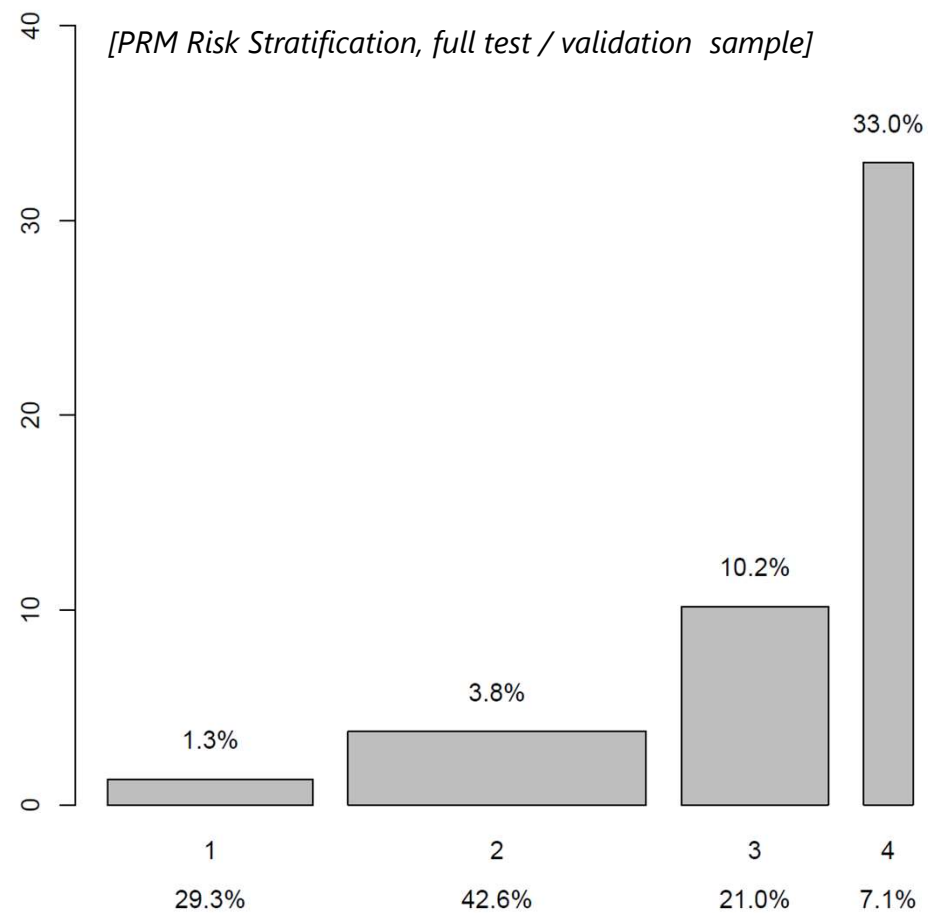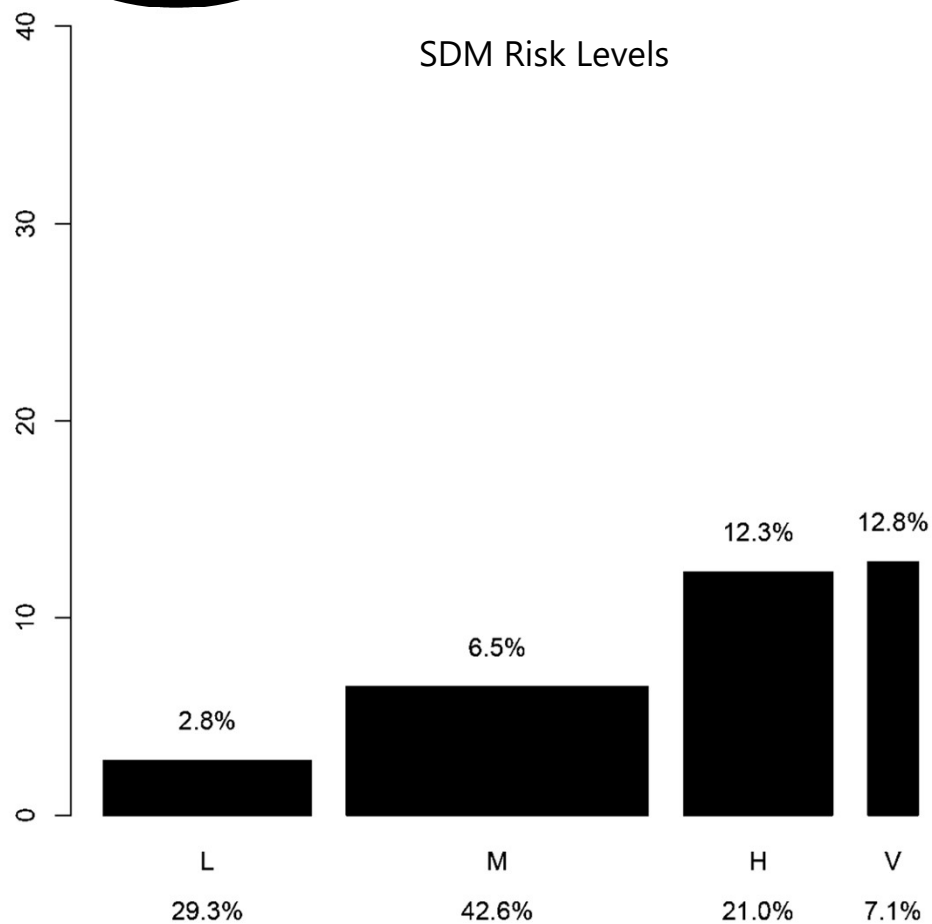
Children designated as "moderate risk" by SDM were 2.5 times as likely to end up in foster care as children classified as "moderate" in risk by the algorithm (4.7% vs. 1.7%).
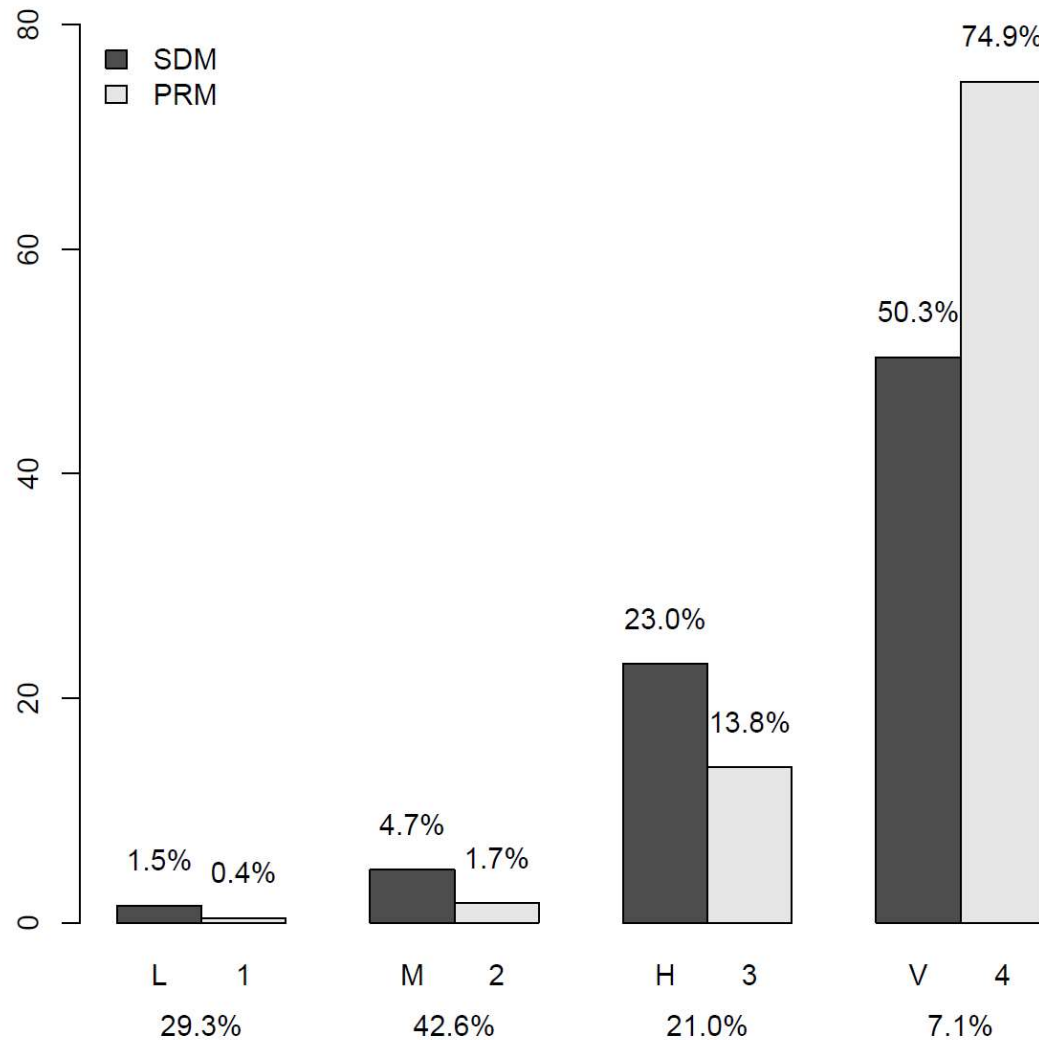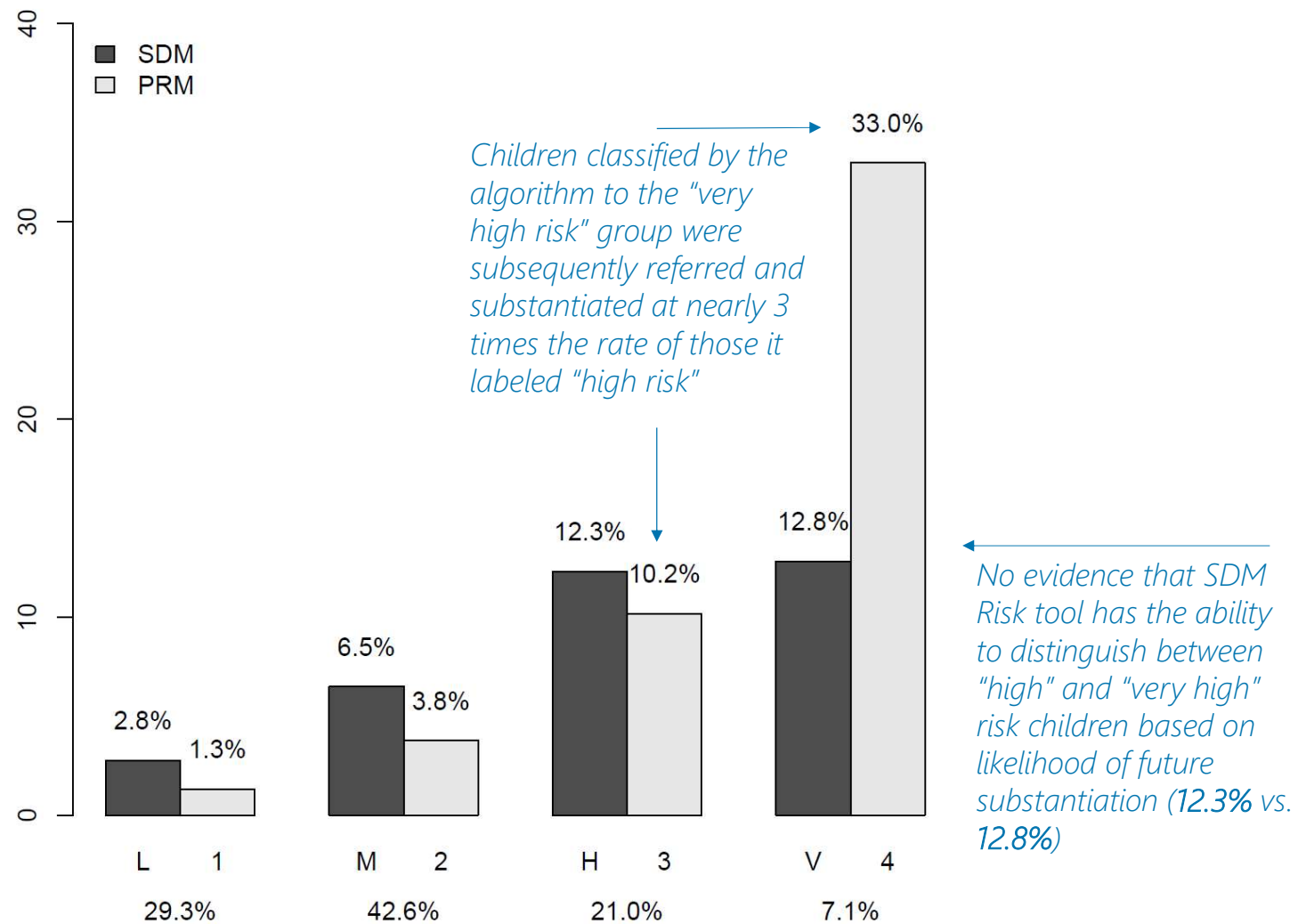

The 7.1% of children identified by the algorithm as "very high risk" had a larger share of children who ended up placed in foster care (74.9%) than did the "very high risk" group identified through the SDM risk assessment tool (50.3%).

Risk stratification based on SDM risk assessment levels vs. PRM stratifications:
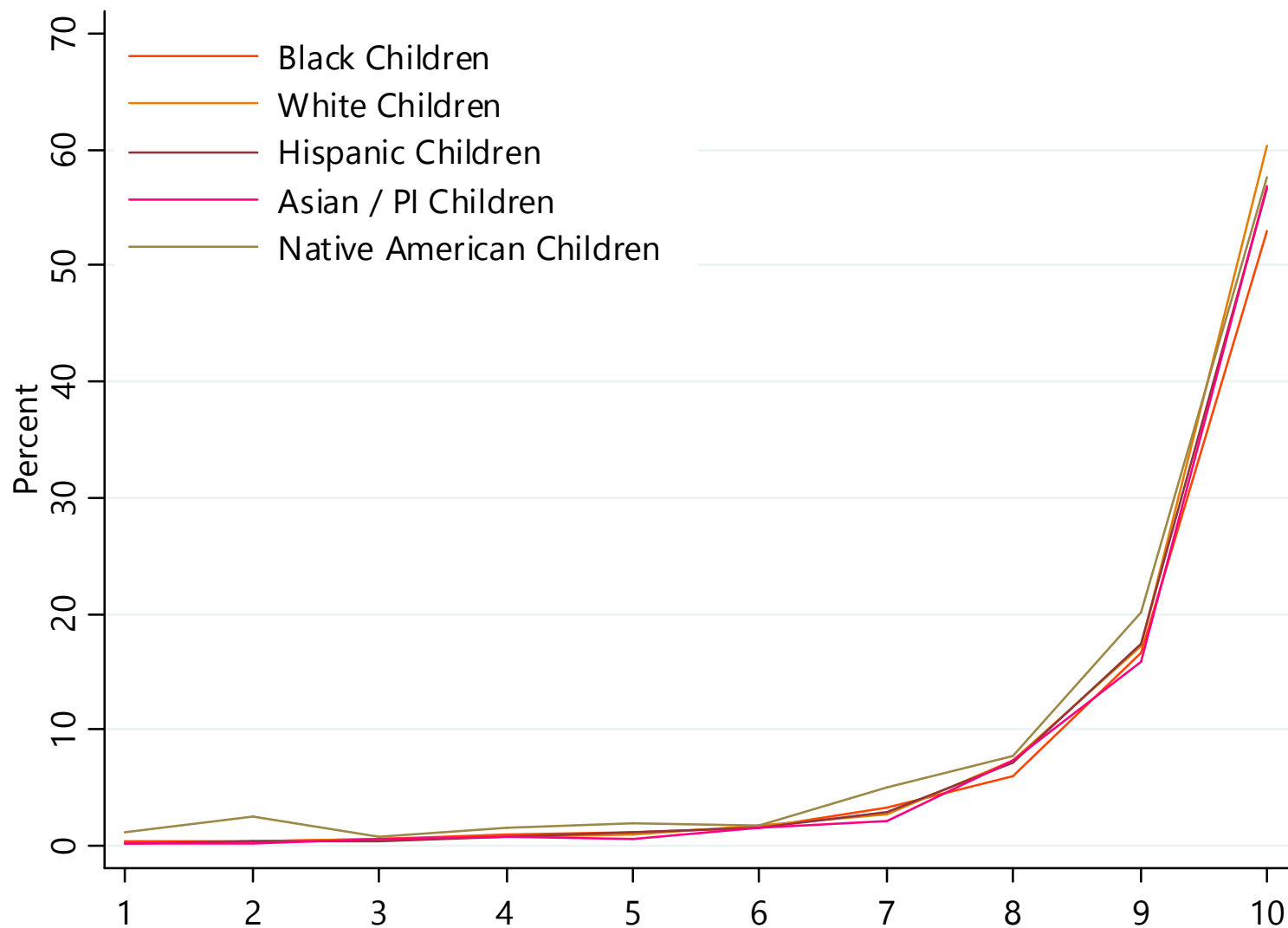*Follow-up substantiated allegation within 24 months*

SDM Risk Levels

[PRM Risk Stratification, full test / validation sample]

	L	M	H	V
	2.8%	6.5%	12.3%	12.8%
	29.3%	42.6%	21.0%	7.1%

	1	2	3	4
	1.3%	3.8%	10.2%	33.0%
	29.3%	42.6%	21.0%	7.1%

substantiation

SDM
PRM

33.0%

Children classified by the algorithm to the "very high risk" group were subsequently referred and substantiated at nearly 3 times the rate of those it labeled "high risk"

12.3%
10.2%
12.8%

6.5%
3.8%

2.8%
1.3%

L   1   M   2   H   3   V   4

29.3%   42.6%   21.0%   7.1%

*No evidence that SDM Risk tool has the ability to distinguish between "high" and "very high" risk children based on likelihood of future substantiation (**12.3%** vs. **12.8%**)*

Equity Checks *[ongoing]*

race / ethnicity
placements

unwarranted variation?

Low Risk Screen-In?

High Risk Screen-Out?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
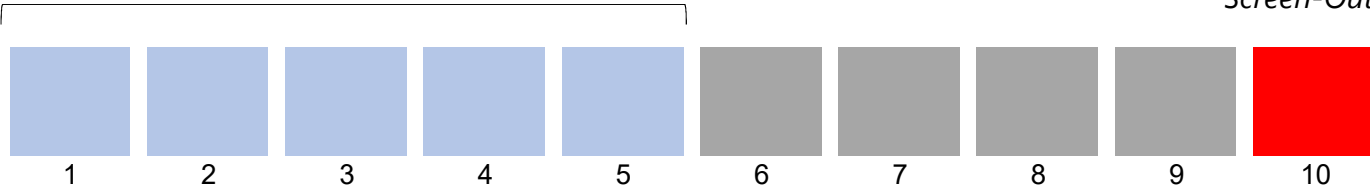
Hispanic Child

Hispanic Child

[67%]

[15%]

Black Child

Black Child

[65%]

[16%]

Low Risk white child/referrals (score of 1-5) are less likely to be screened-in (59%) than low risk black and Hispanic child/referrals (similarly scored as 1-5s).

High risk white child/referrals (score of 10) are more likely to be screened out (21%) than high risk black and Hispanic child/referrals (similarly scored as 10s).

Although relatively small percentage point differences, thousands of families would be affected by even modest shifts to greater equity in screening practices by race.

White Child

White Child
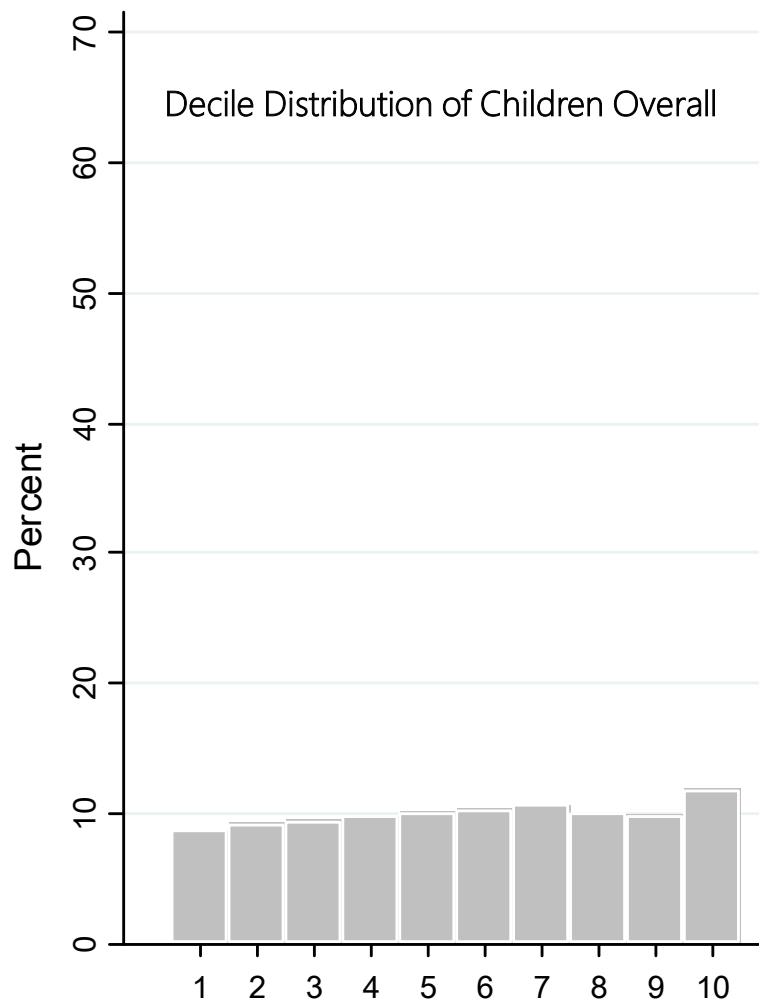
[59%]

[21%]

8 percentage point difference, highest to lowest

6 percentage point difference, highest to lowest

External Validation *[ongoing]*

Decile Distribution of Children Overall

Decile Distribution of Children Near-Fatally or Fatally Maltreated

Maltreatment Near-Fatalities & Fatalities among children with a history of maltreatment reports

*[**maximum** decile score]*

Deciles 9-10
78%

59%

19%

Deciles 1-5
<10%

Percent

1 2 3 4 5 6 7 8 9 10
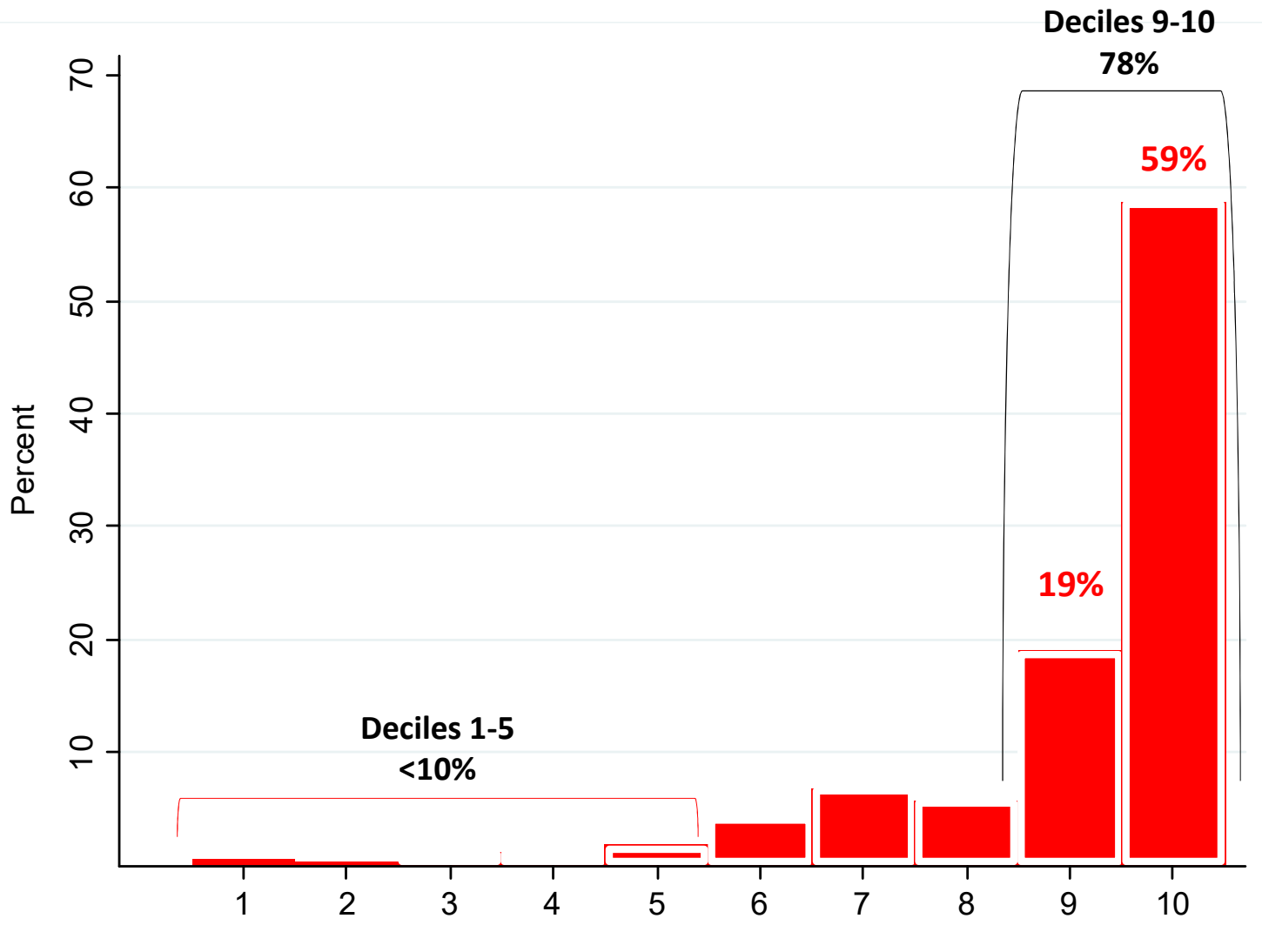
*Among children who experienced a near-fatality or fatality due to maltreatment, **59%** would have been scored in the top decile at the time of at least one referral, **78%** would have fallen in the top two deciles.*

Maltreatment Near-Fatalities & Fatalities among children with a history of maltreatment reports

*[**maximum** decile score]*
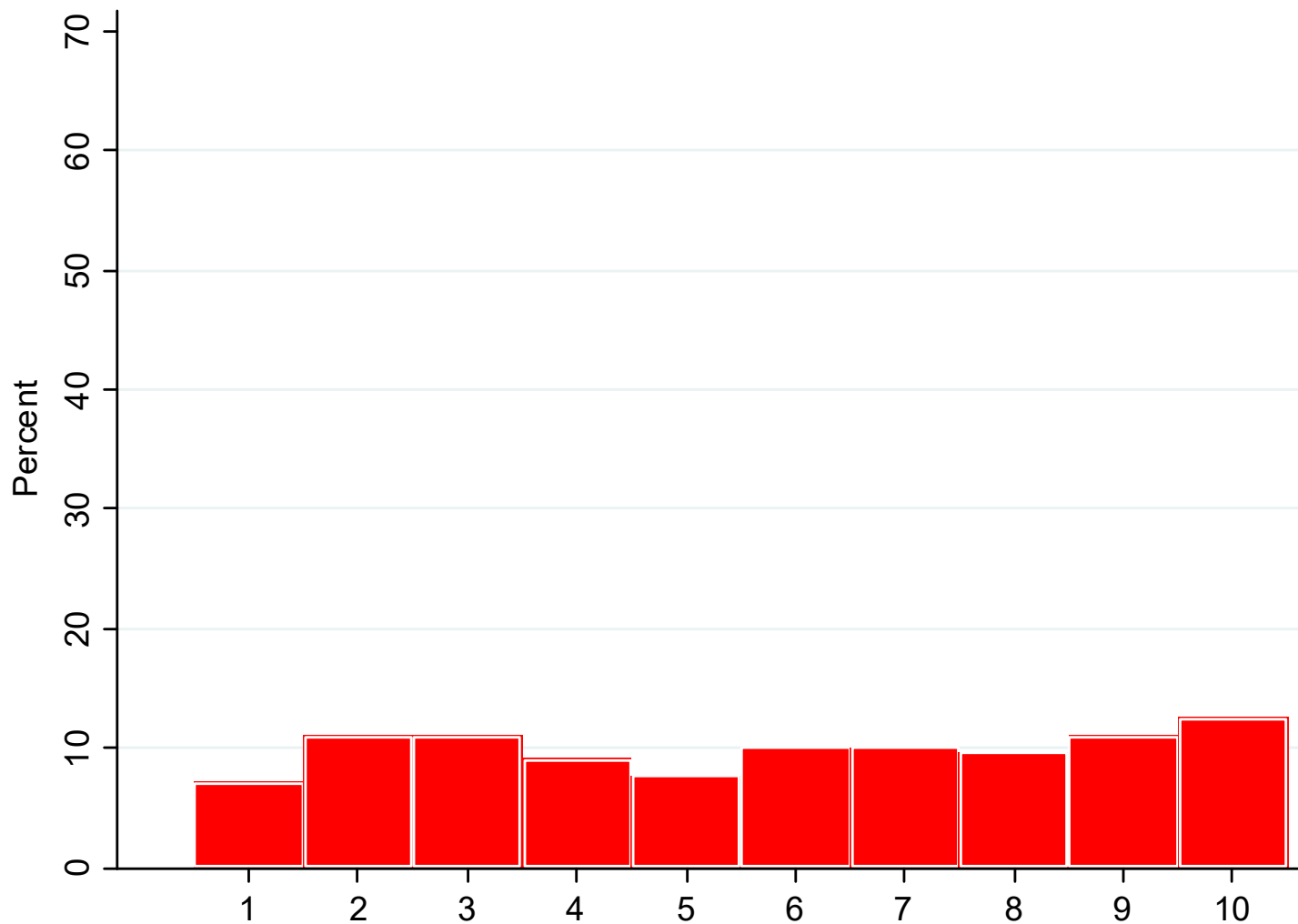
Cancer deaths among children with a history of maltreatment reports? (n=207)

Among children who die from cancer and who have had at least one referral for maltreatment, no relationship with risk score (as expected).

team

Research Team

- Emily Putnam-Hornstein, USC
- Rhema Vaithianathan, AUT
- John Prindle, USC
- Stephanie Cuccaro-Alamin, USC/UCB
- Huy Nghiem, USC
- Tanya Gupta, USC *(now, Amazon Lab126)*

Project Advisors / Consultants

- Diana Benavides Prado, AUT
- Alexandra Chouldechova, CMU
- Erin Dalton, Allegheny County
- Jacquelyn McCroskey, USC
- Michael Mitchell, USC
- Barbara Needell, Consultant
- Daniel Webster, UCB

Ethical Reviewers

- Brett Drake, WashU
- Melissa Jonson-Reid, WashU

State Data & Research Partners

- CDSS, Office of Child Abuse Prevention
- CDSS, Research Services Branch
- CDSS, Child Protection and Family Support Branch
- Child Welfare Digital Services

County Data & Practice Partners

- Los Angeles County Department of Children & Families
- Monterey County Department of Social Services
- San Francisco Human Services Agency

University Collaborators

- AUT Centre for Social Data Analytics
- UCB California Child Welfare Indicators Project

[please note that the findings presented and conclusions drawn emerge from the research team and do not necessarily reflect the views of other partners in this work]
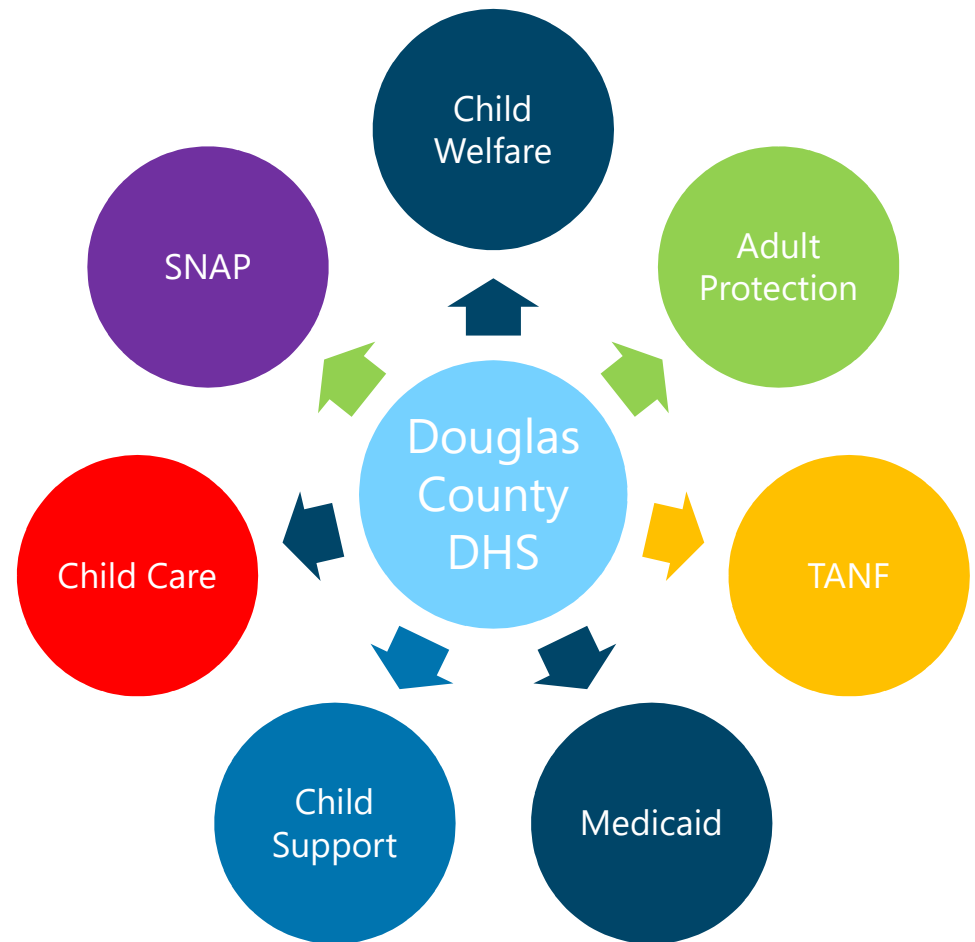
The Douglas County Experience

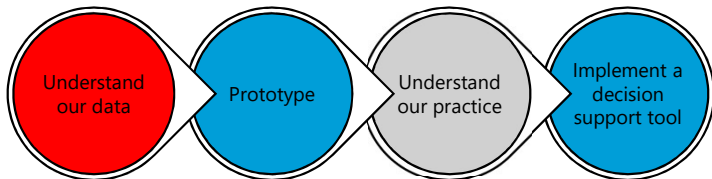Developing the Douglas County Decision Aid

Douglas County

Population:	**328,632**
Median Income:	**$109,292.00**
Child Population:	**89,388**
Race/Ethnicity:	**83.2% Caucasian**
High School Graduate:	**98.48%**
Bachelor's Degree :	**59.43%**
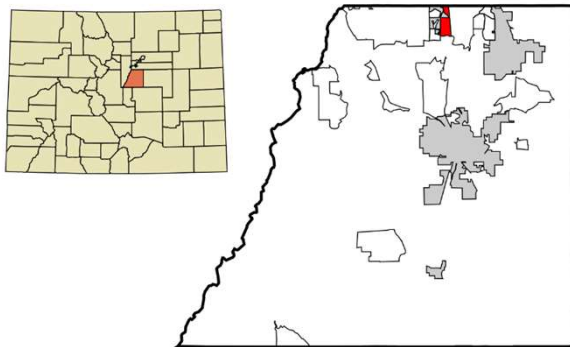
Douglas County Human Services

Child Welfare

Adult Protection

SNAP

Douglas County DHS

TANF

Child Care

Child Support

Medicaid

Douglas County Data

Trails- Child Welfare Case Management

- Basic Referral Level Information
- Client Level Information
- Referral Outcome/Disposition
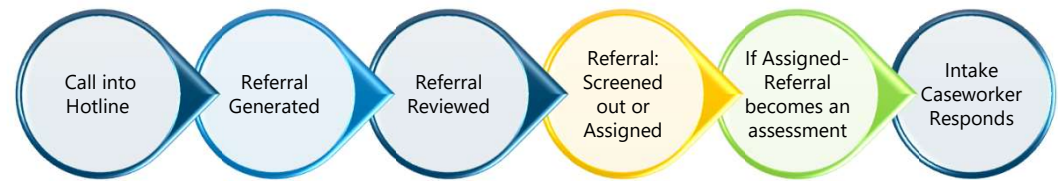- Assessment Information

CBMS- Public Benefits System

- Client Level Information
- Eligibility/Benefit Program Participation

Douglas County HSCARES

- Acts as Douglas County's Master Client Index
- Allows tracking of households and services provided

The Problem

Referral Assignment Decisions

Currently Process Through Red Team

Call into Hotline → Referral Generated → Referral Reviewed → Referral: Screened out or Assigned → If Assigned- Referral becomes an assessment → Intake Caseworker Responds
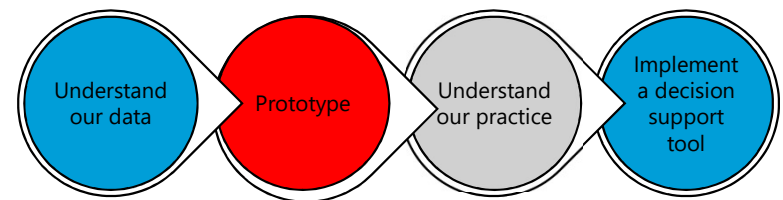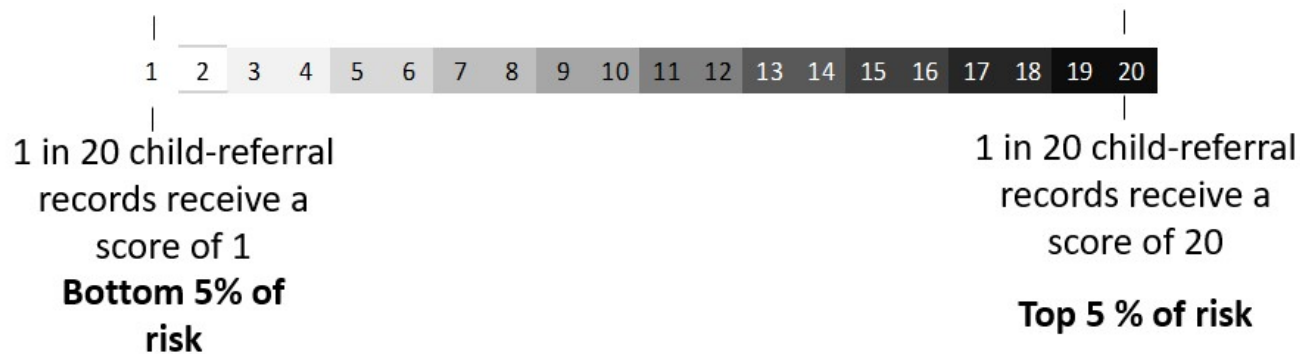
<u>R</u>eview <u>E</u>valuate <u>D</u>irect

2 Caseworkers and a Supervisor

Analyze Current Referral
All Child Welfare History
Criminal History for all parties
CBMS (Food Stamp, Medicaid)

Have to make a VERY difficult decision that is SUBJECTIVE!
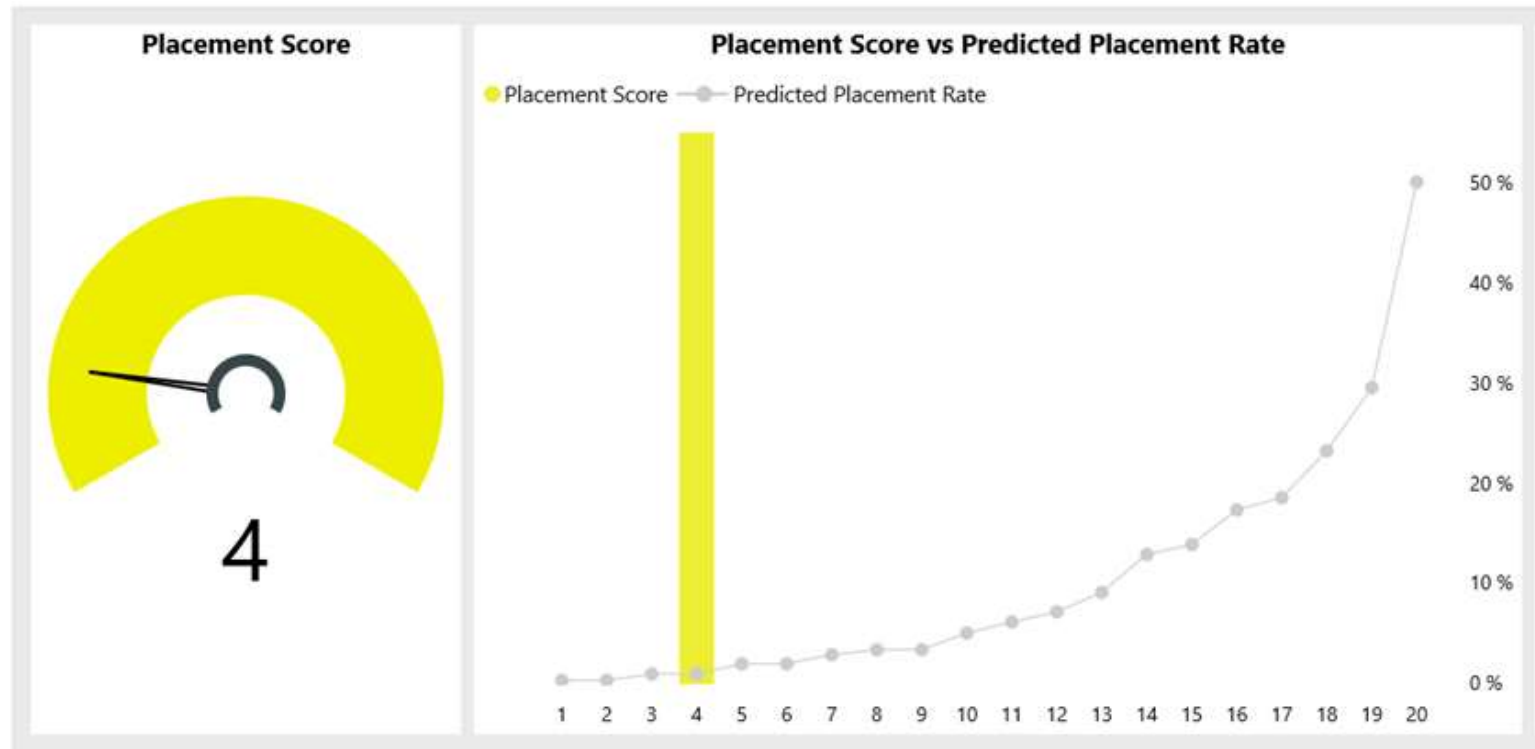
Generating the Scores

- Each child on the referral receives a predicted risk score of 1 to 20

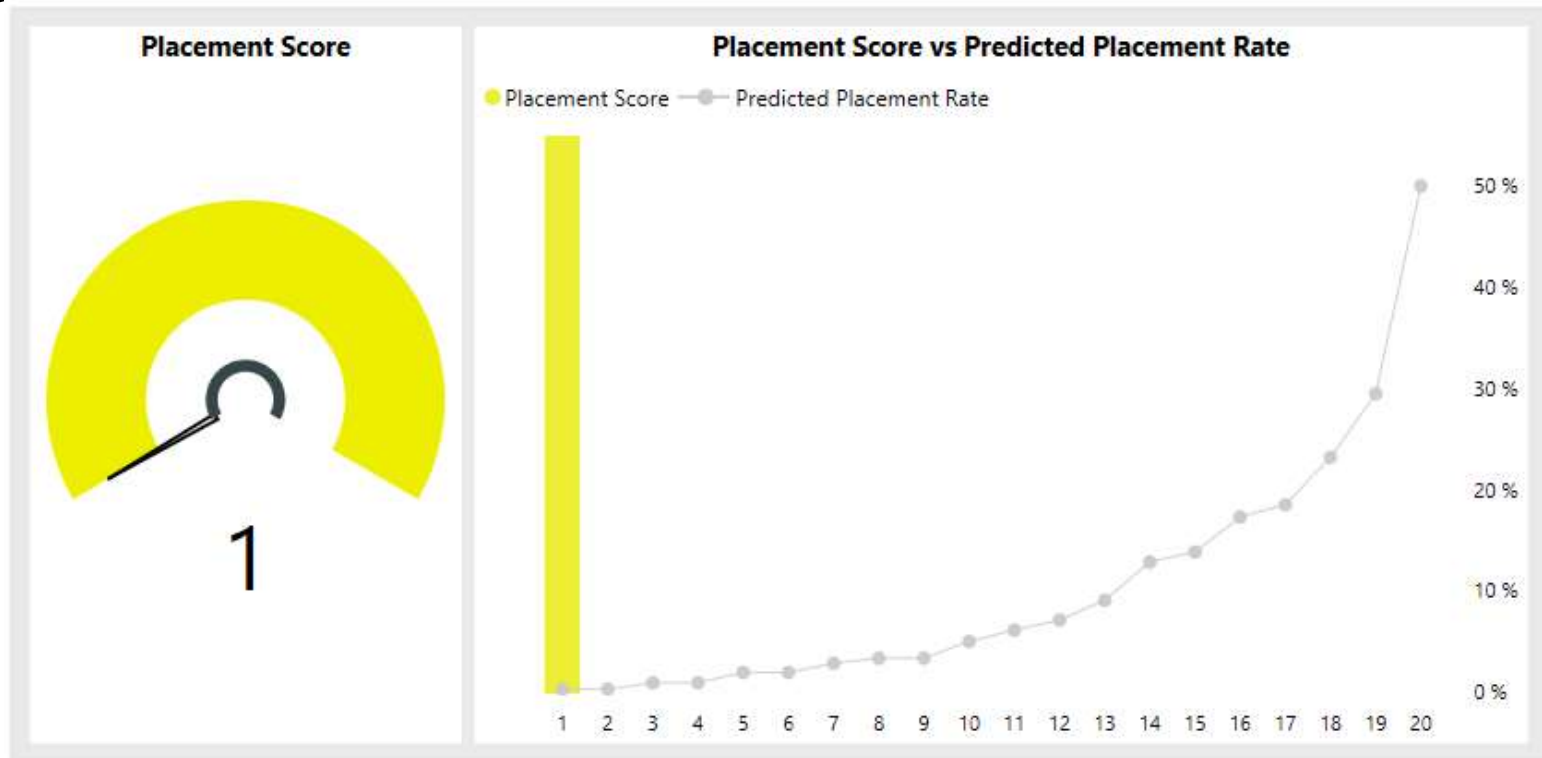- The higher the score, the higher the chance of re-referral or placement

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |

1 in 20 child-referral records receive a score of 1
Bottom 5% of risk

1 in 20 child-referral records receive a score of 20
Top 5 % of risk

Understand our data — Prototype — Understand our practice — Implement a decision support tool

Carrie

Original RED Team
Decision: ████████

Placement Score

4

Placement Score vs Predicted Placement Rate

● Placement Score ──○── Predicted Placement Rate

Cynthia B

Original RED Team
Decision:

Monique

Original RED Team
Decision: ▓▓▓▓▓▓▓▓
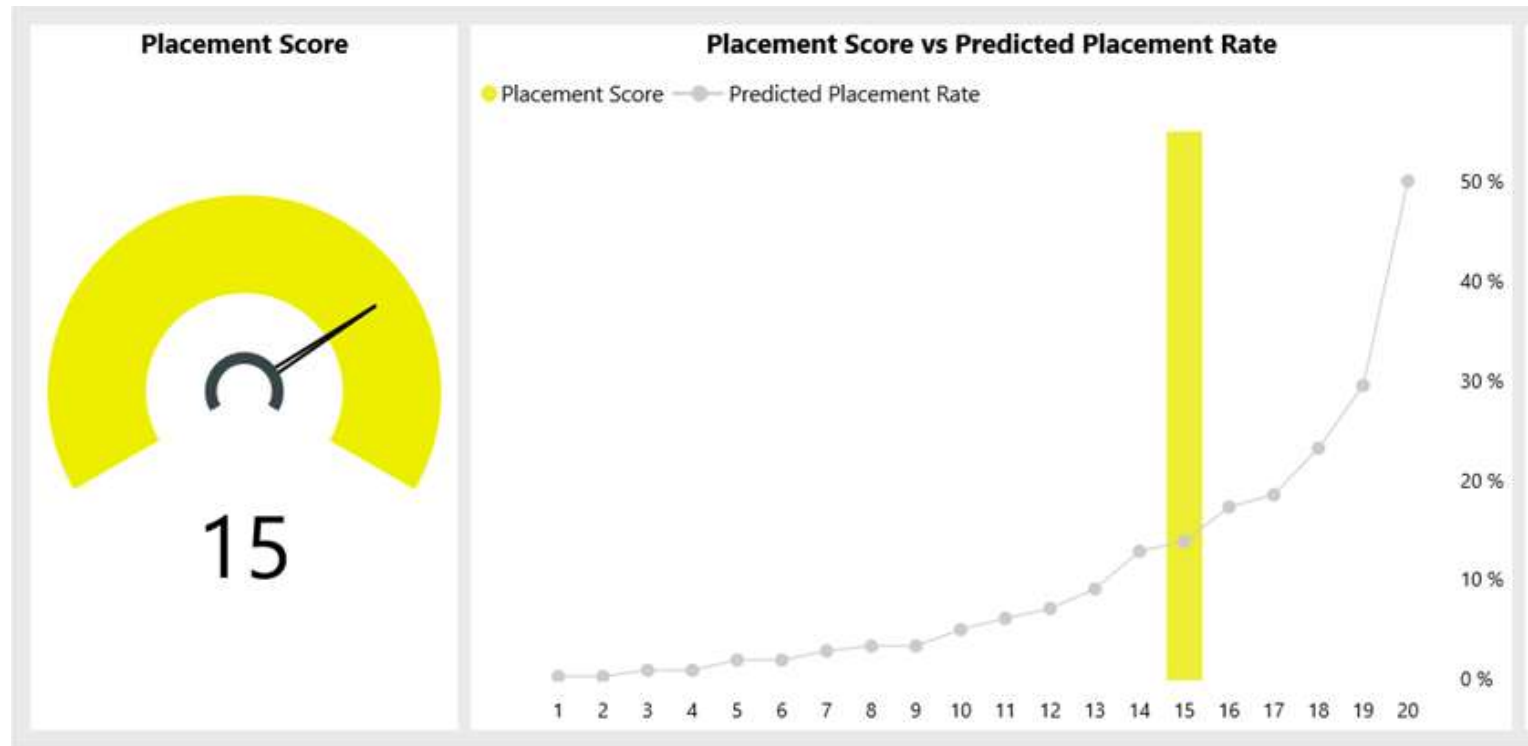
Naomi

Original RED Team
Decision: ▮▮▮▮▮▮▮

Susan

Original RED Team
Decision: ████████

Placement Score

1

Placement Score vs Predicted Placement Rate

● Placement Score ── Predicted Placement Rate
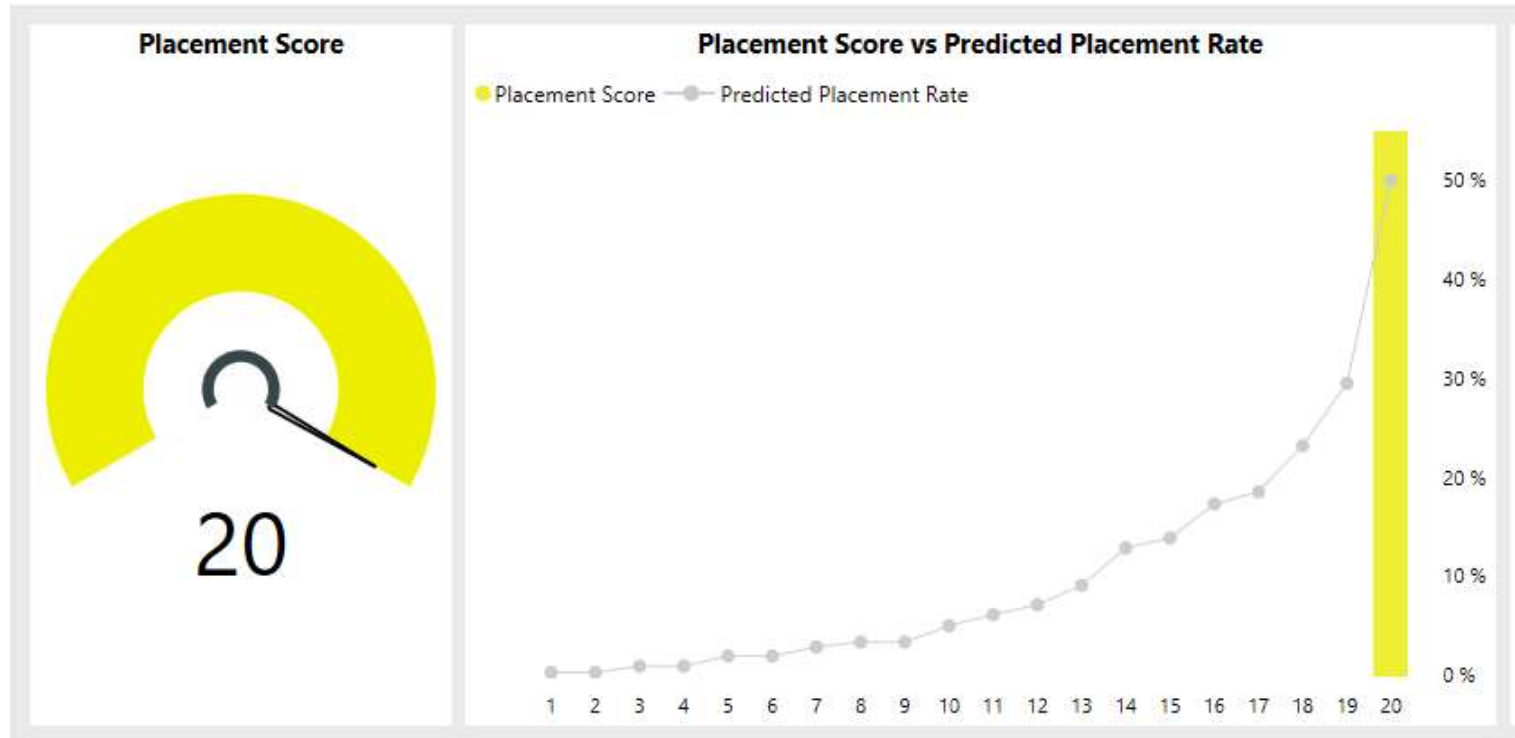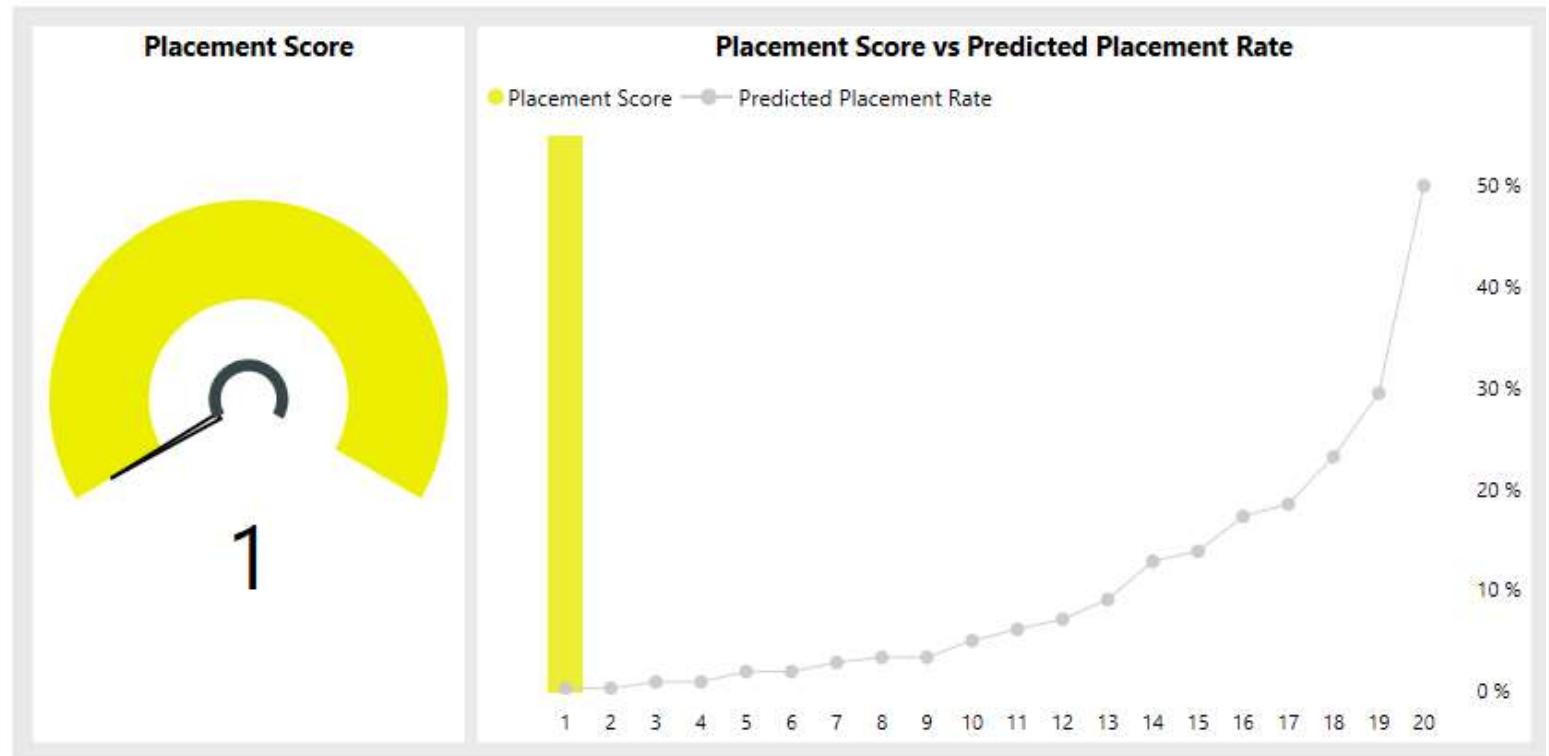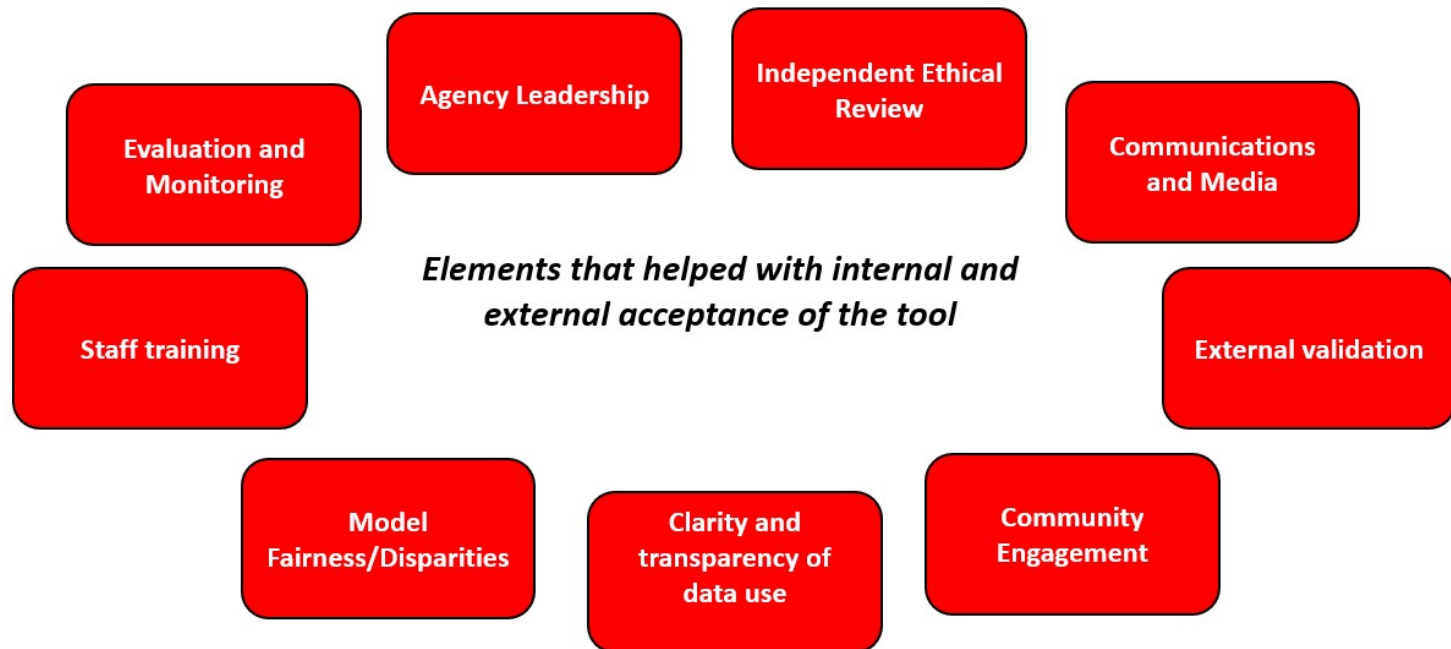
Reflections from Douglas County

- ➢ **Value of insights into practice**

- ➢ **Experience working with research team**

- ➢ **Important lessons**

- ➢ **Staff Response to screening score and tool**

How can a predictive risk tool be used to determine candidacy?

Elements of Success

Agency Leadership

Independent Ethical Review

Evaluation and Monitoring

Communications and Media

Elements that helped with internal and external acceptance of the tool

Staff training

External validation

Model Fairness/Disparities

Clarity and transparency of data use

Community Engagement

Funding provided by,

Thank You

Other Partners

Thank You

Questions?

Emily Putnam-Hornstein   ehornste@usc.edu

Dan Makelky   dmakelky@douglas.co.us

Ruby Richards   rarichar@douglas.co.us

Nicole Adams   nadams@douglas.co.us