



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

May 8, 2019

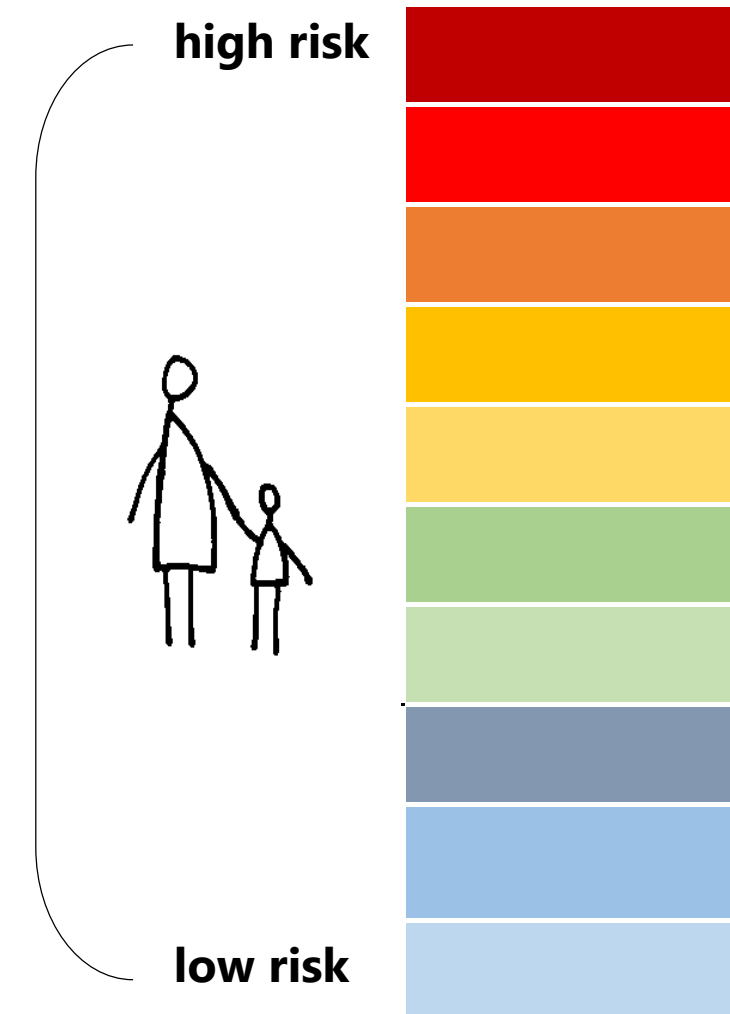


**prediction is very difficult,
especially if it's about the future**

Niels Bohr

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*)



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

- Resistance from clinicians and other frontline staff
- "Black Box"
- Risk does not mean willingness to accept intervention
- Only as good as intervention delivered
- Bias that may be "baked" into data

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

Home » American Journal of Public Health (AJPH) » February 2017

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

Accepted: October 24, 2016 Published Online: January 11, 2017

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.



One might conceptualize child welfare agencies as social service agencies, but that would be incorrect. In reality, child welfare agencies are gate-keepers and the workers decision makers.

- Gelles & Kim, 2008

Allegheny Family Screening Tool



Screening ScoreHistorical 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.

Last Run By:
Jane McBeth

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

Algorithm Versions Used
Re-referral v43
Placement v22

High

Moderate

Low

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



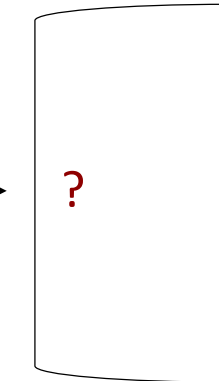
Completion of SDM
Hotline Tool



Worker Supervision



Screen-In



Screen-Out

Red Team
Conferencing

?

Worker
Assignment

?

Specialized
Unit

?

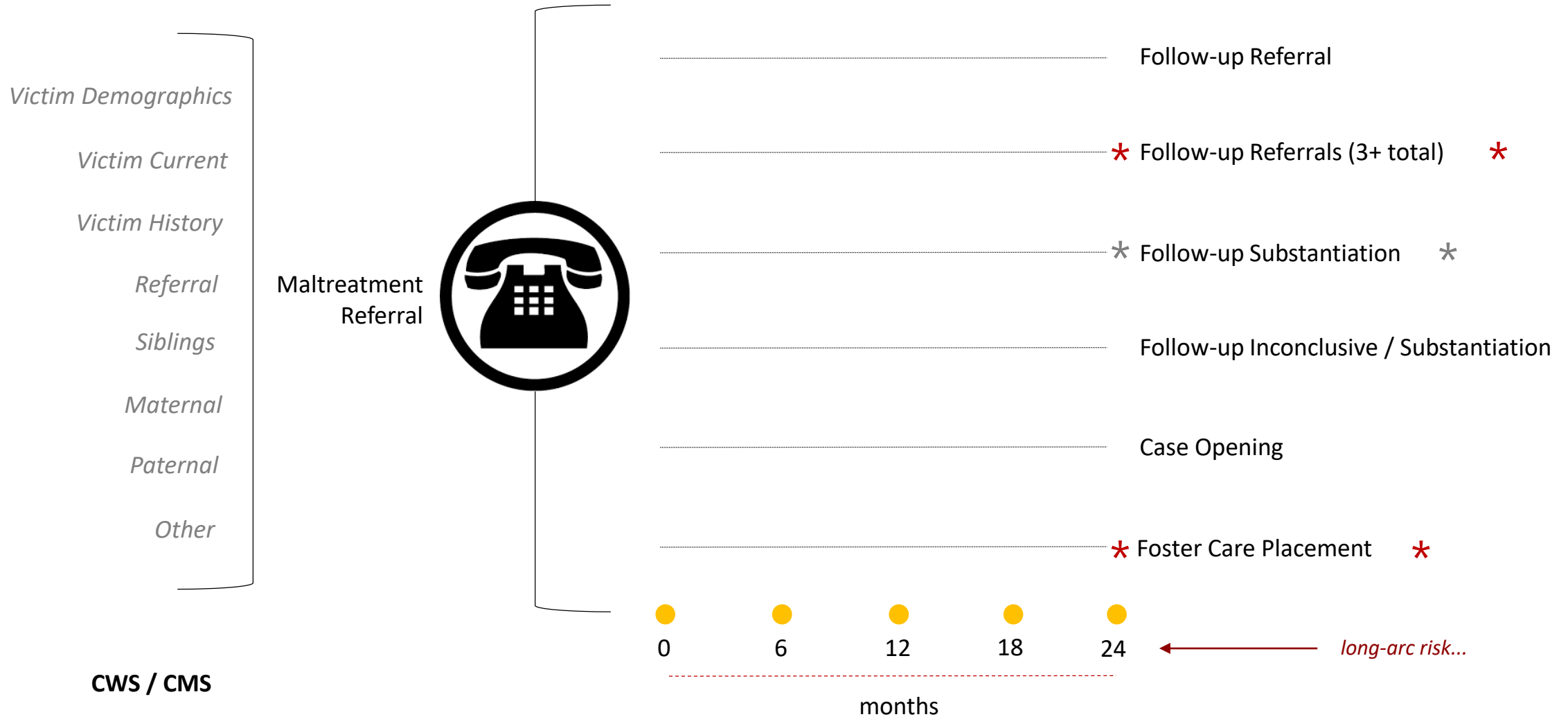
Level of
Supervisory
Oversight

?

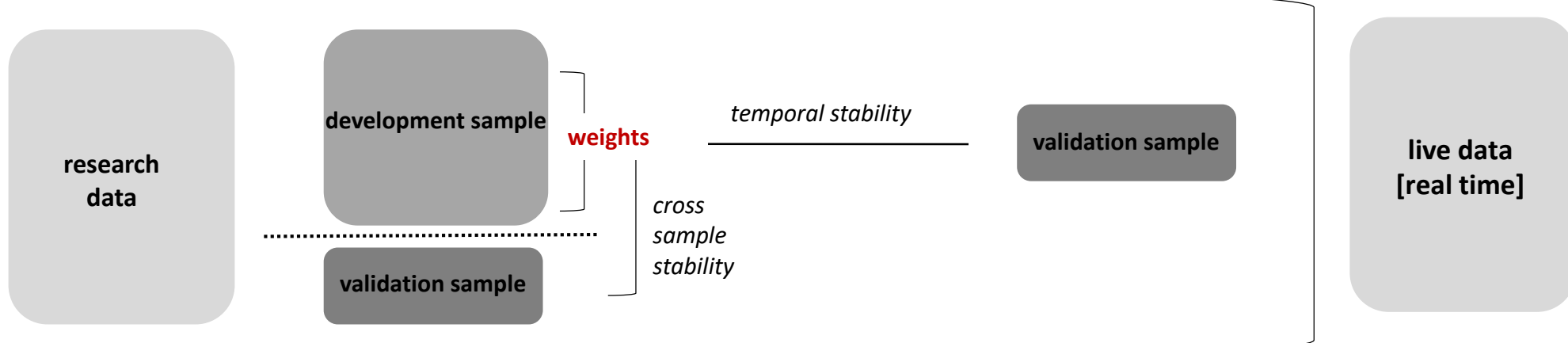
Candidacy for
FFPSA

?

[system] OUTCOMES



ALGORITHM DEVELOPMENT & VALIDATION



Data

3.4 million CWS/CMS observations (referral * child) from 2010-2014
1.2 million unique children used for training / testing model
300+ CWS/CMS predictors coded, plus SDM hotline fields

Development Sample

Repeated random samples drawn to develop initial predictor weights for model
Final predictor weights based on model trained using data for 920,000 unique children / unique families

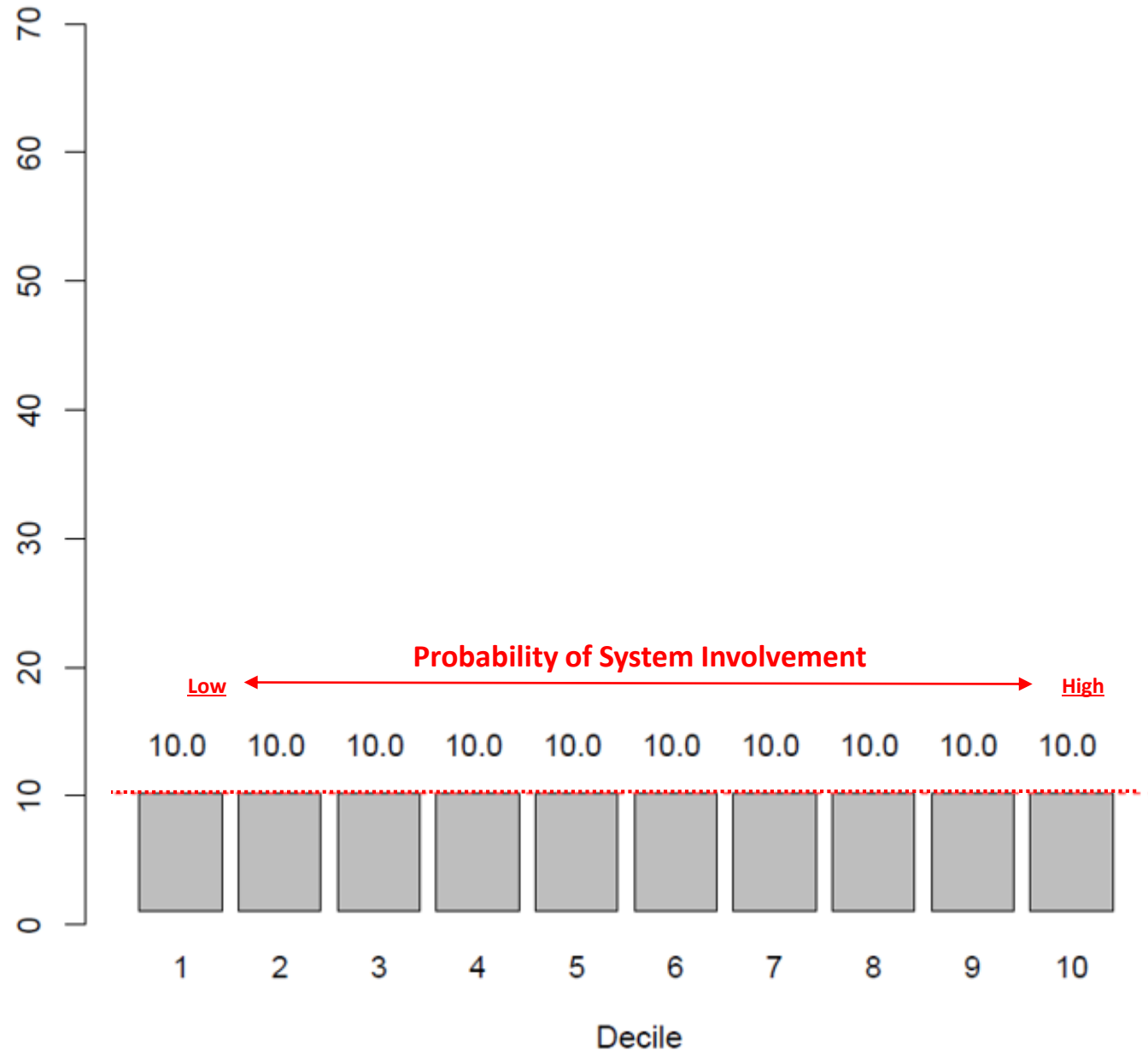
Validation Sample

Repeated random samples (records not used in model development) used to validate / test initial accuracy of weights
Final model estimates validated for 230,000 unique children / unique families (80 / 20 split)

County-specific validations

"risk deciles"

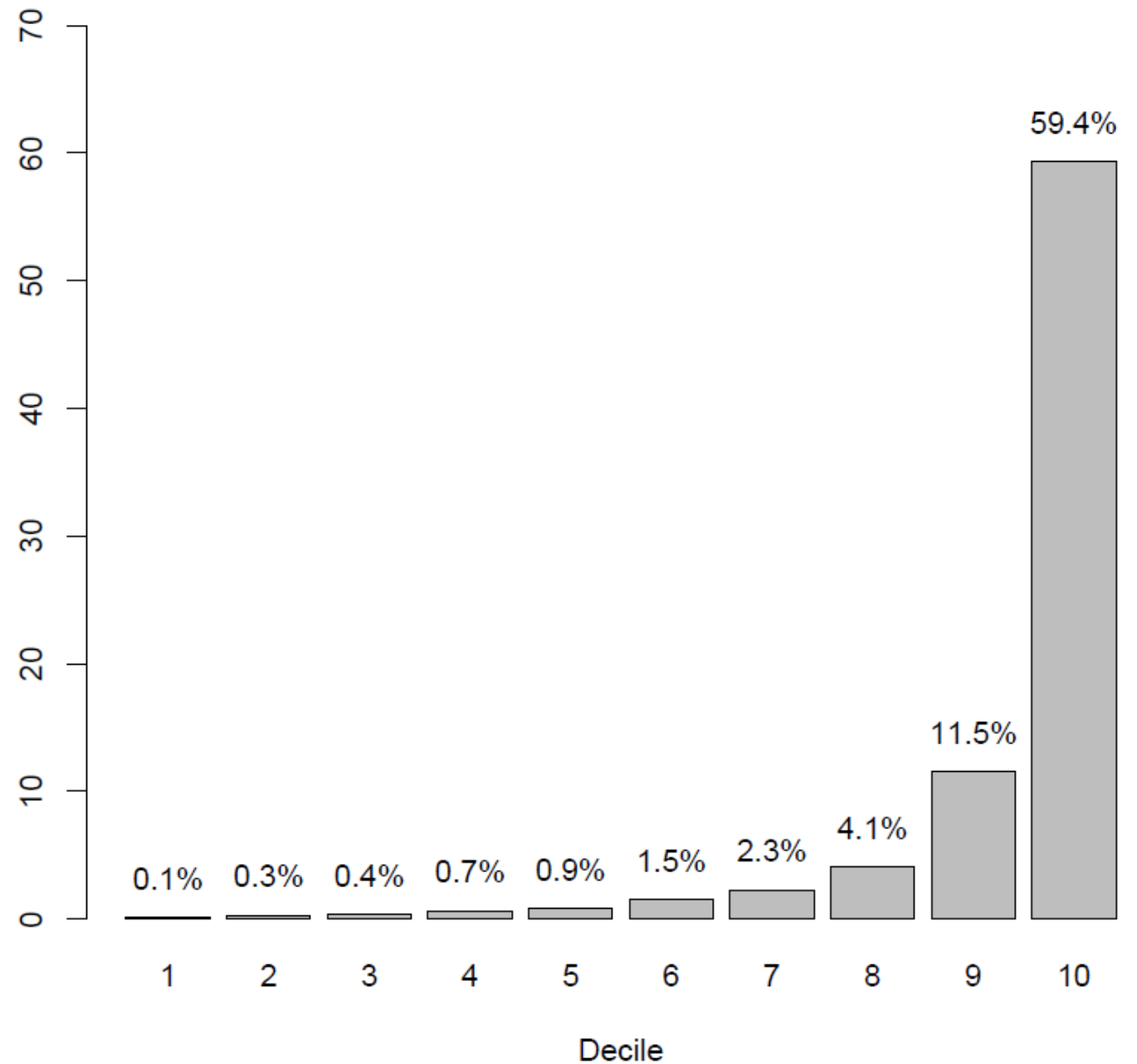
We use the algorithm to assign each child/referral into 10 evenly sized groups (or deciles) based on the predicted probability of system involvement.



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%)



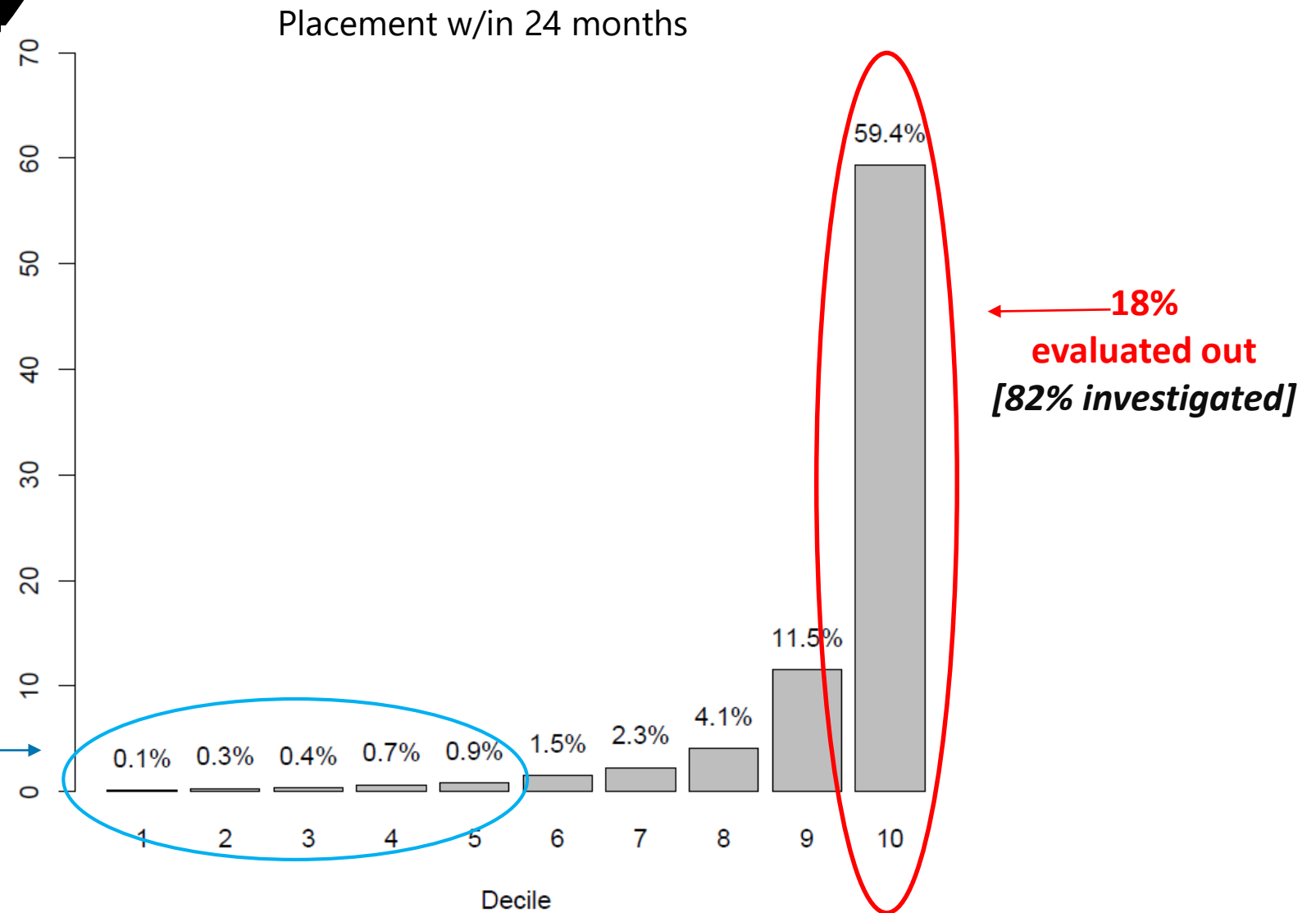
Value Add..

"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.

37%
evaluated out
[63% investigated]



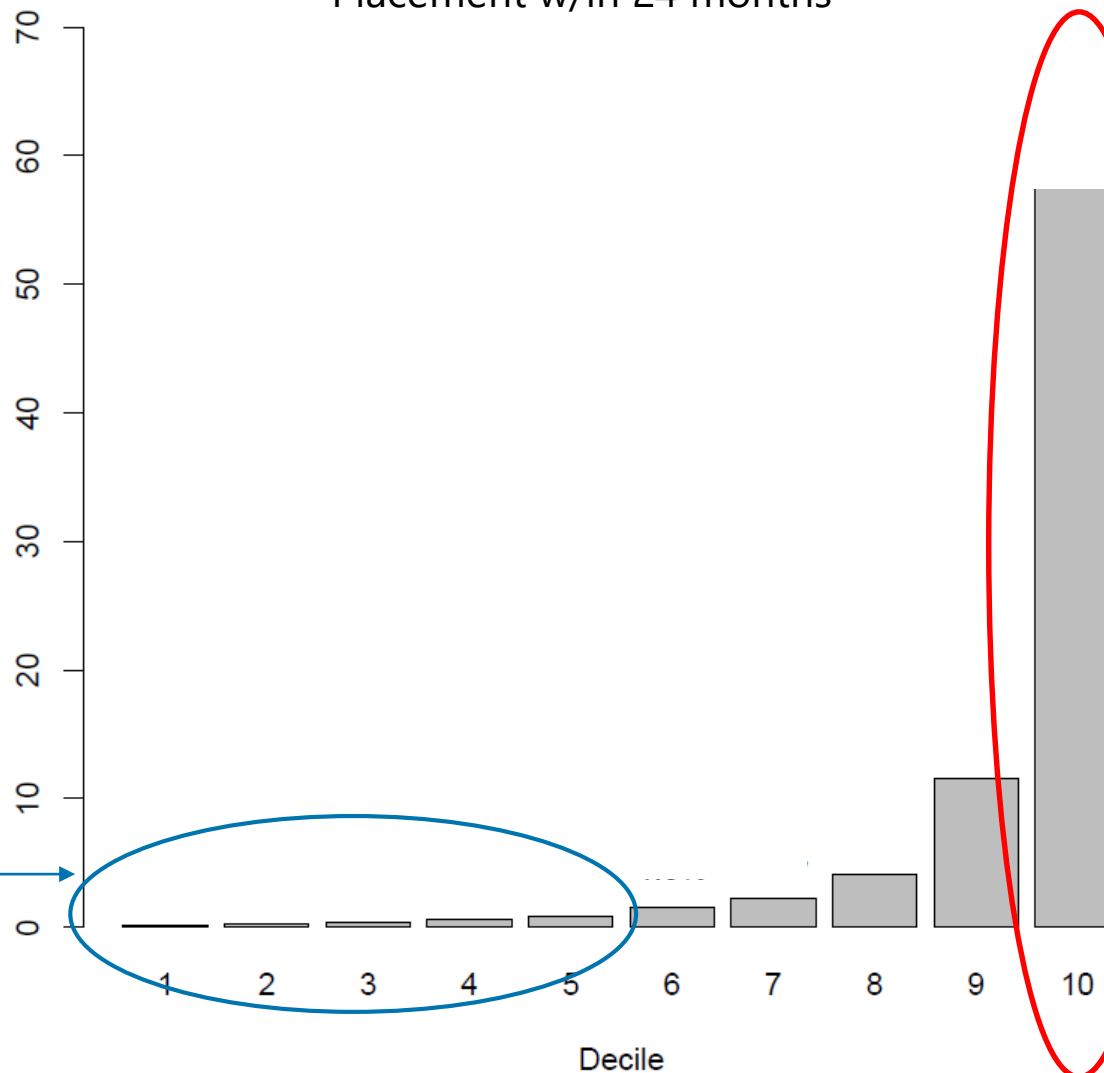
placement
outcomes?

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**

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**

SDM® family risk assessment



"The SDM family risk assessment identifies families with low, moderate, high, or very high probabilities of future abuse or neglect. By completing the risk assessment, the worker obtains an objective appraisal of the likelihood that a family will maltreat their child in the next 18 to 24 months. 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."

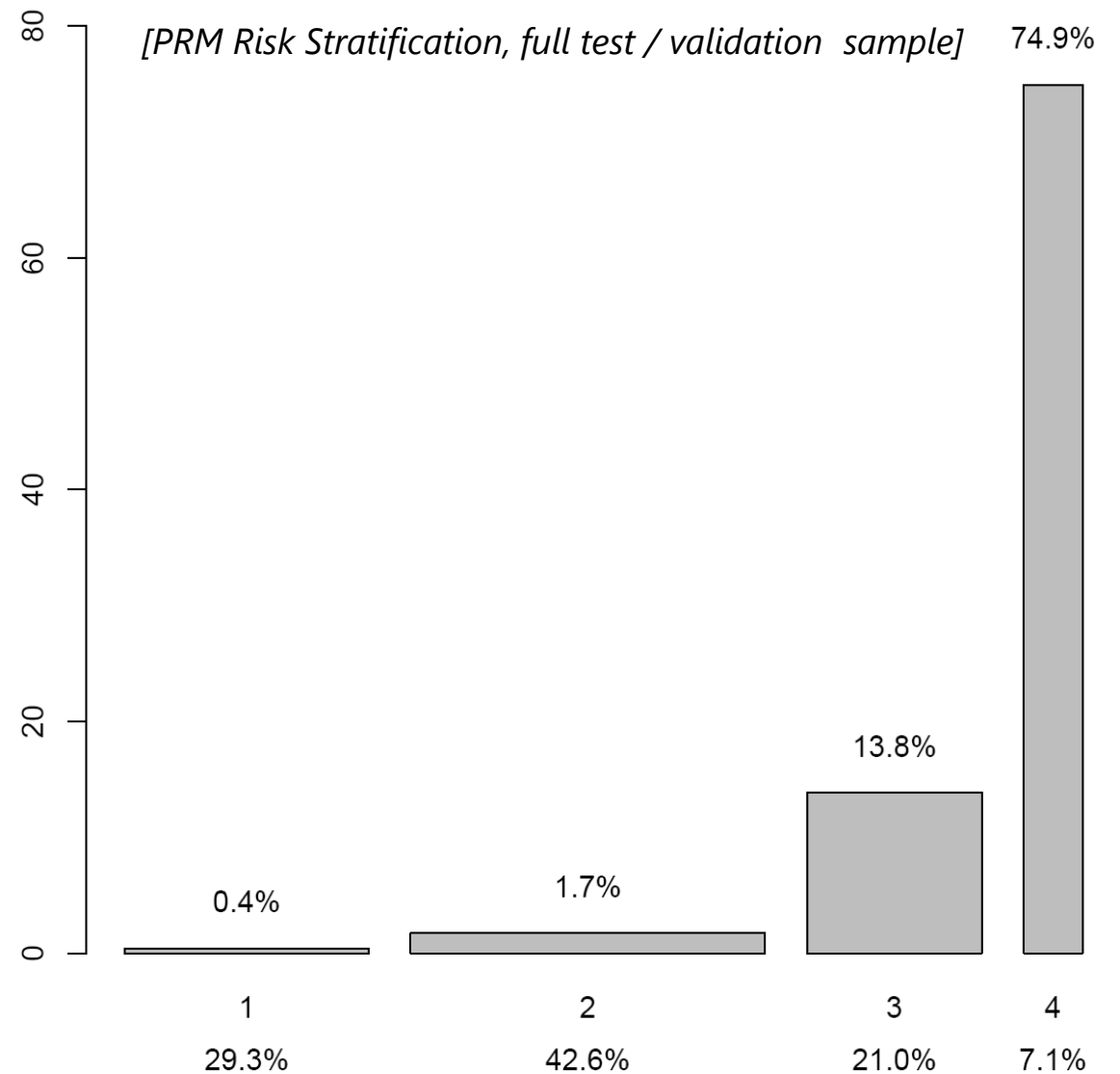
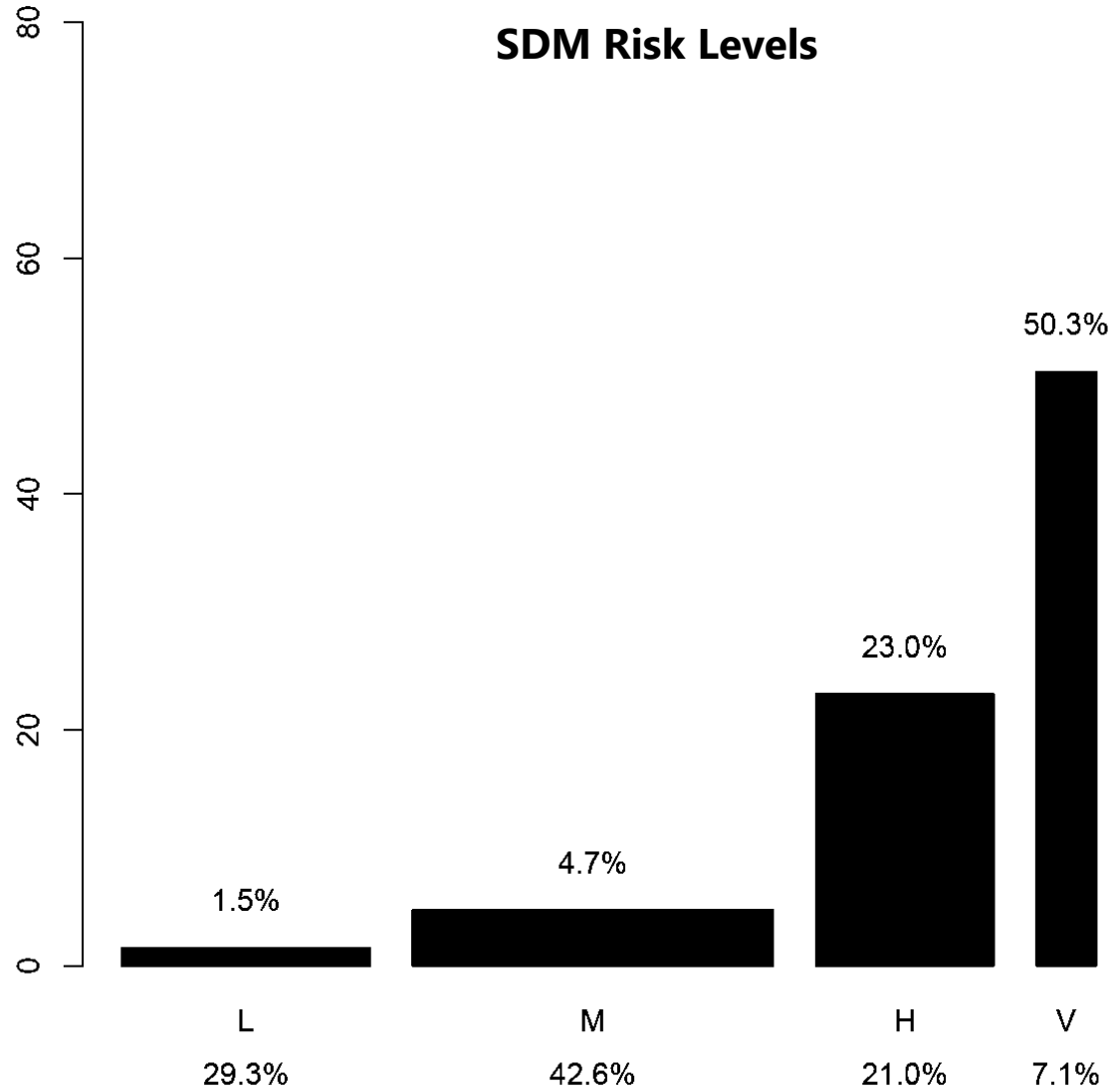
http://www.childsworld.ca.gov/res/pdf/SDM_Manual.pdf

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

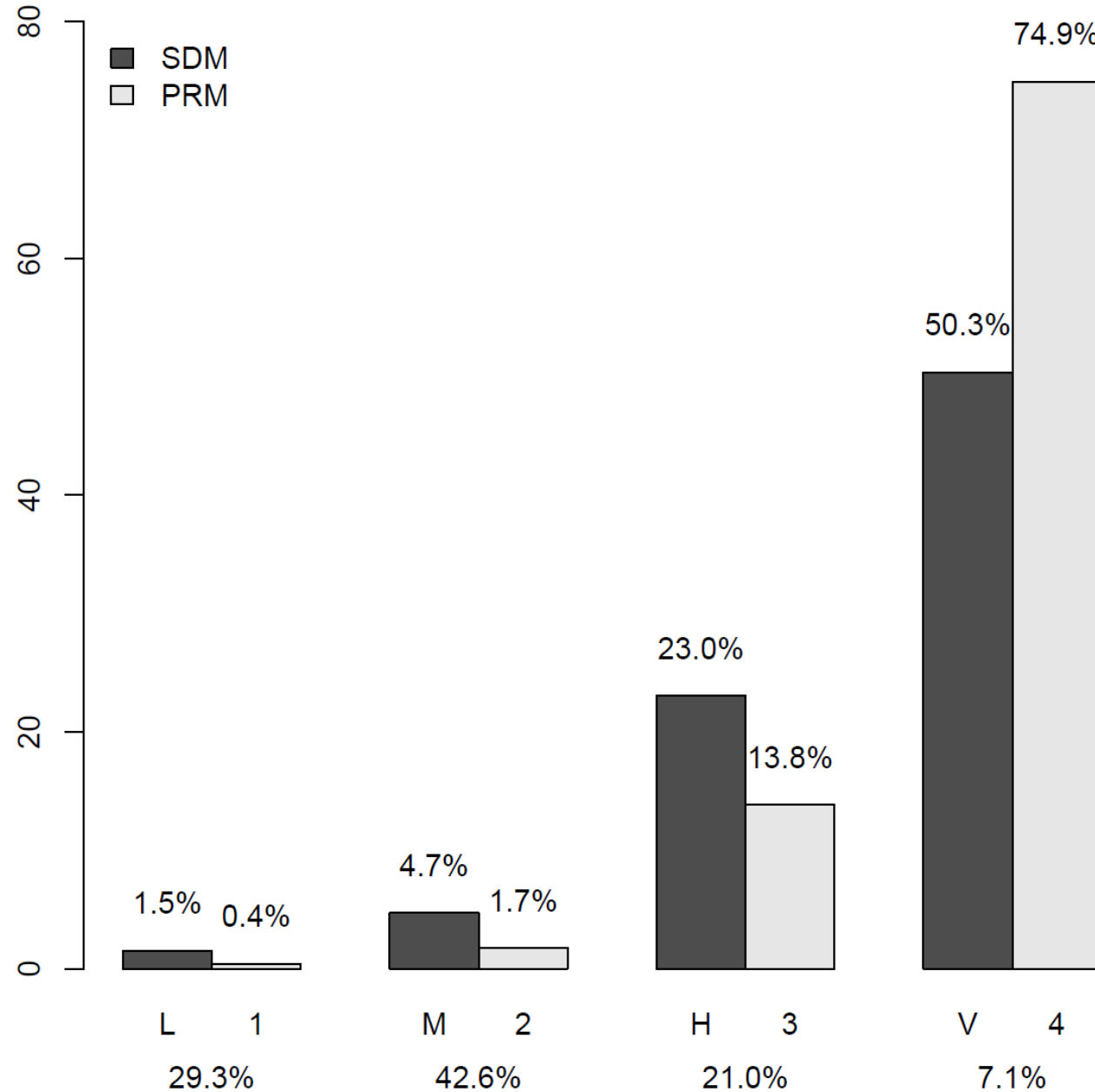
Risk stratification based on SDM risk assessment levels vs. PRM stratifications:
Placements in Foster Care within 24 months



placement

Current use of the SDM risk assessment results in nearly half (42.6%) of children classified as "moderate risk".

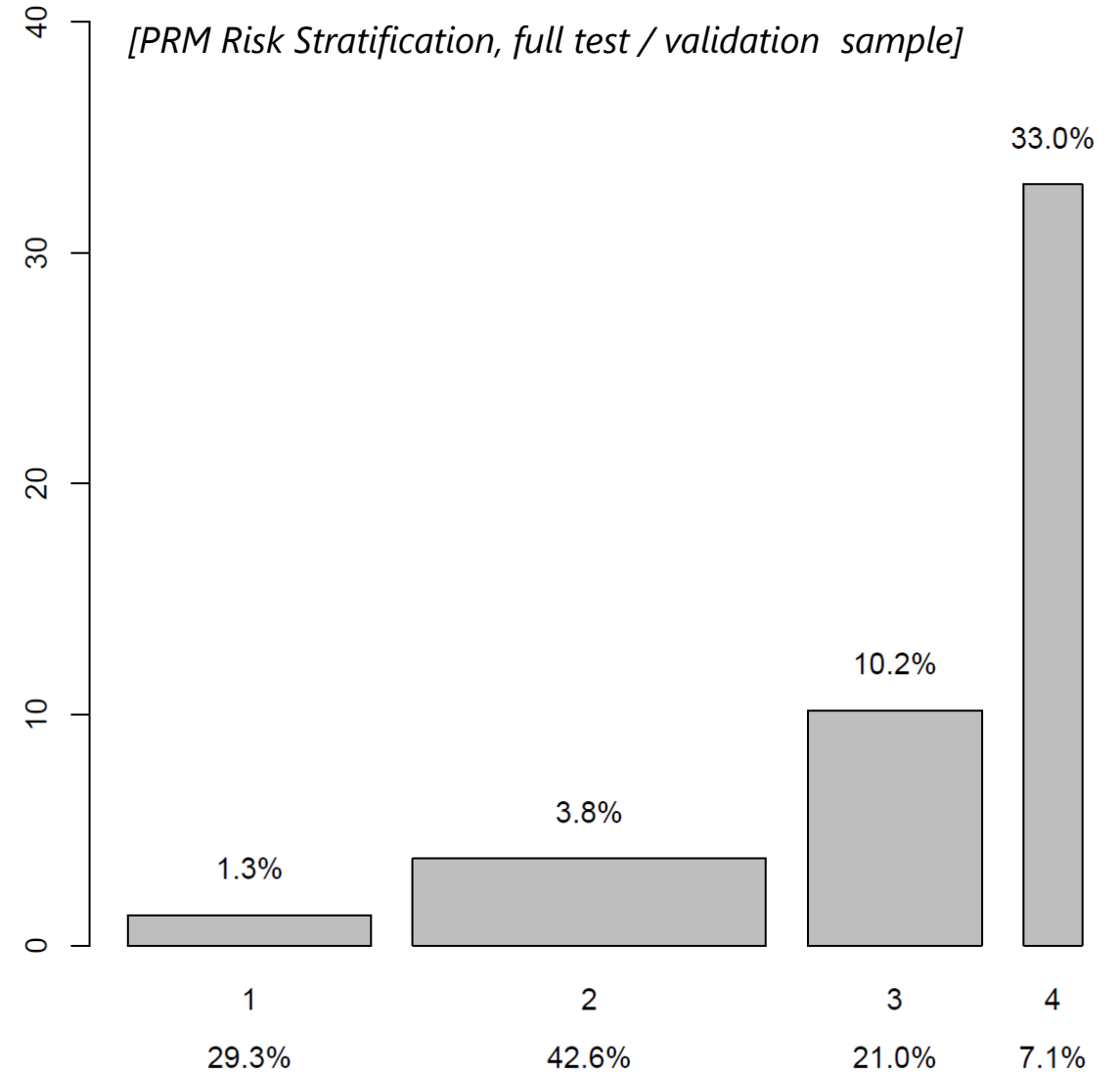
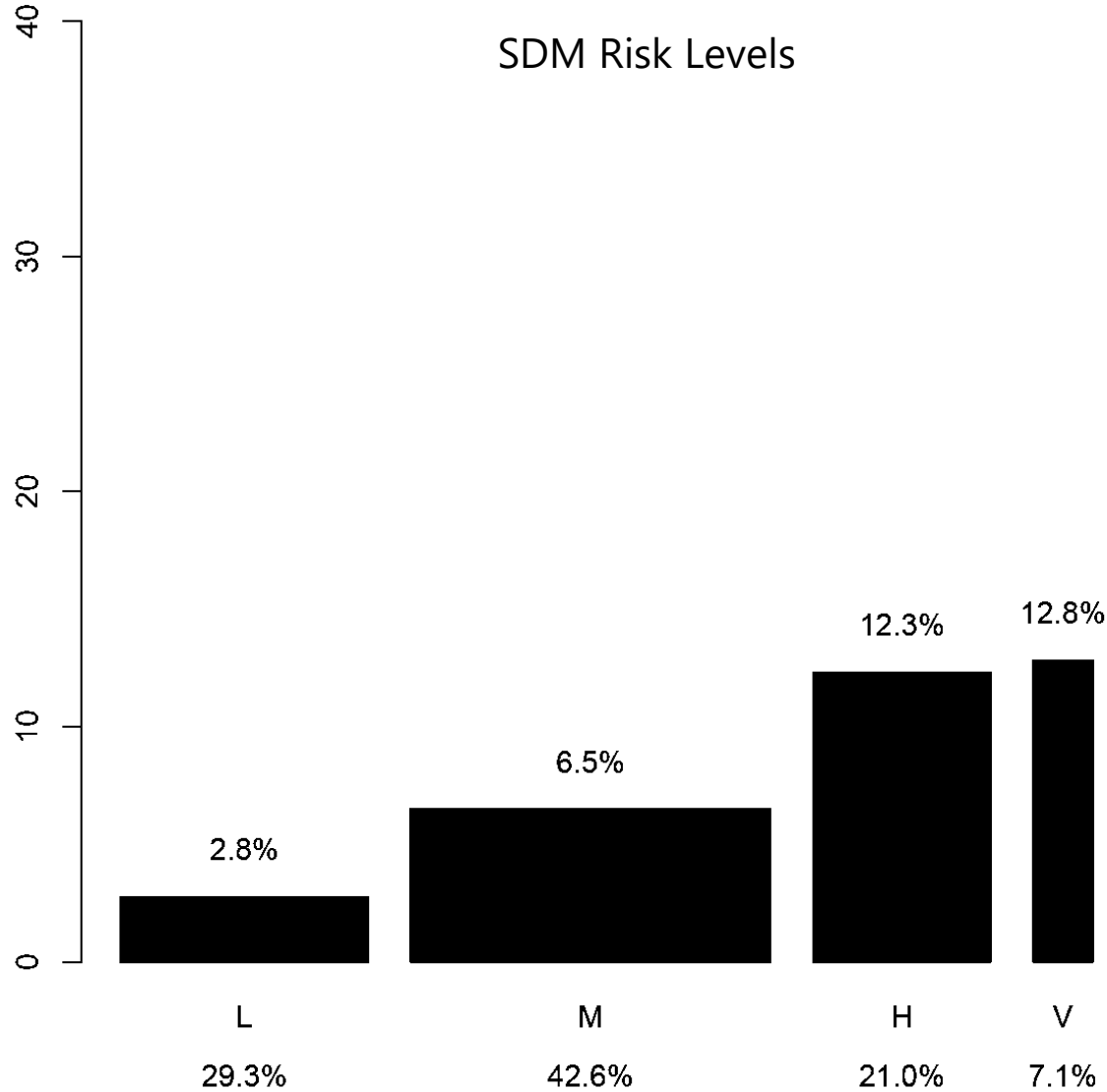
Children designated as "moderate risk" by SDM were 2.5 times as likely to end up in foster care as children classified as "moderate" in risk by the algorithm (4.7% vs. 1.7%).



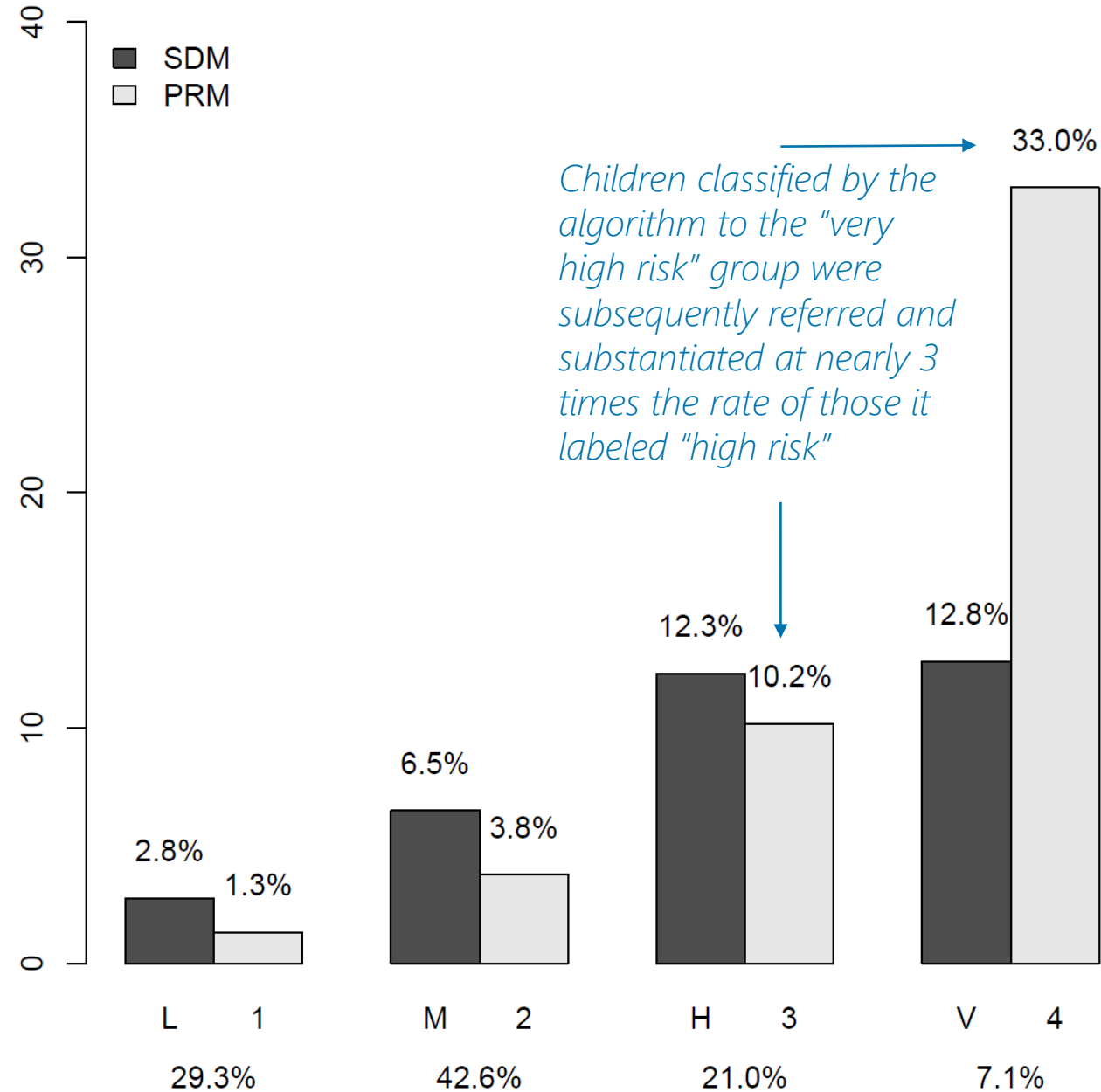
← The 7.1% of children identified by the algorithm as "very high risk" had a larger share of children who ended up placed in foster care (74.9%) than did the "very high risk" group identified through the SDM risk assessment tool (50.3%).

substantiation

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



substantiation

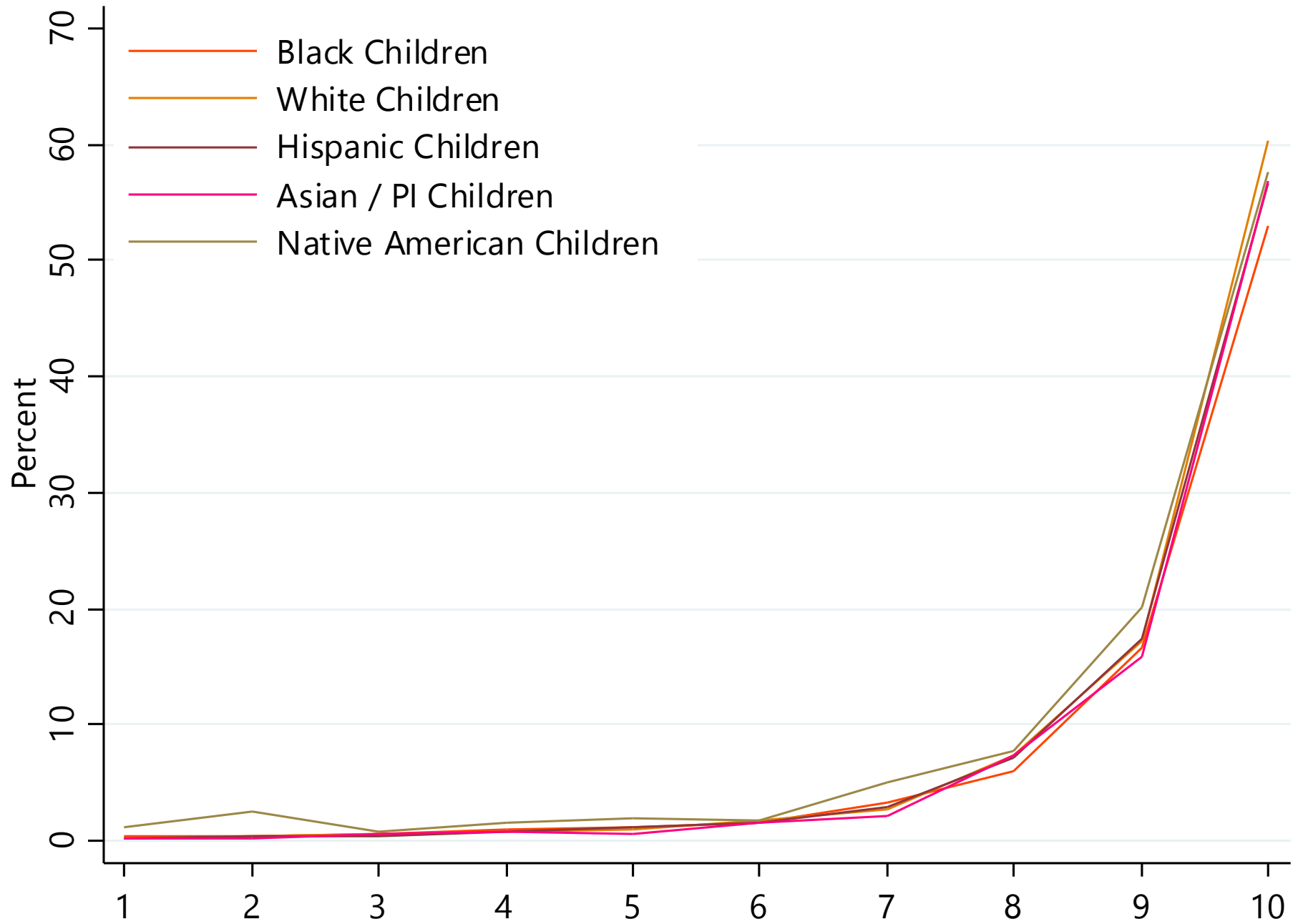


Children classified by the algorithm to the "very high risk" group were subsequently referred and substantiated at nearly 3 times the rate of those it labeled "high risk"

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?



Hispanic Child

Hispanic Child



[67%]

Black Child



[15%]

Black Child



[65%]

White Child



[16%]

White Child



[59%]



[21%]

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).

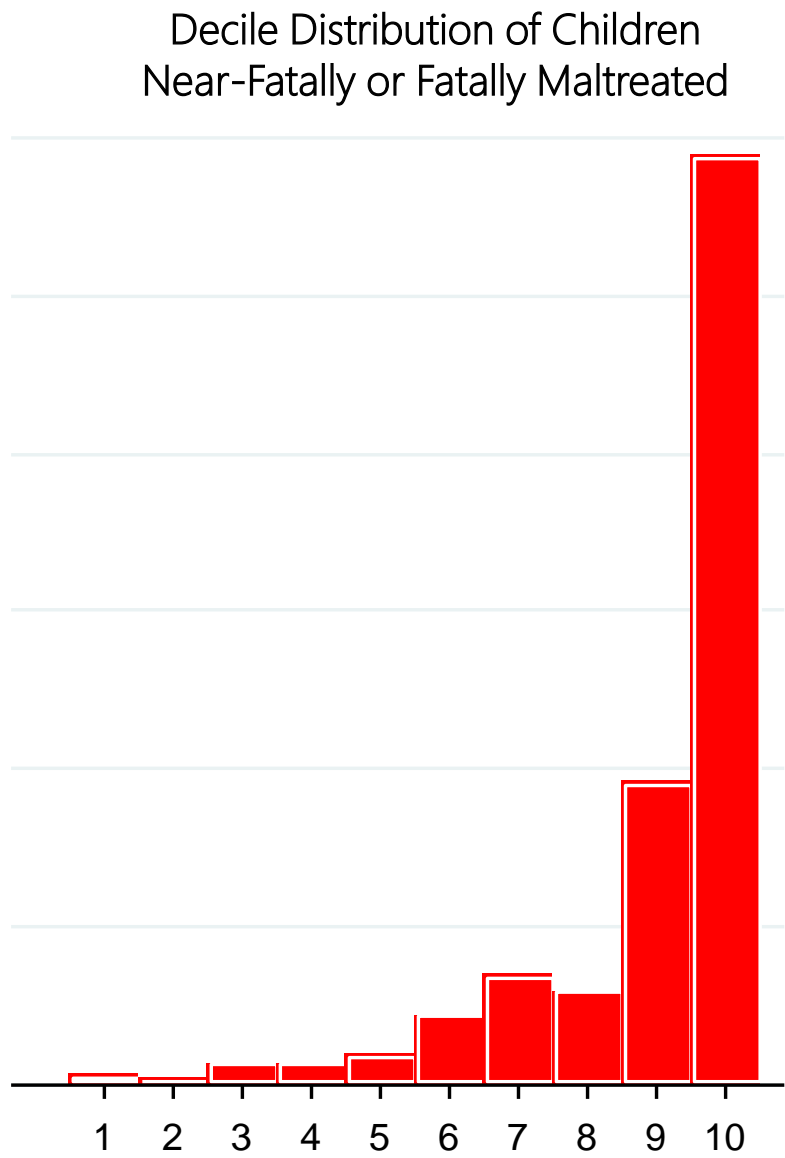
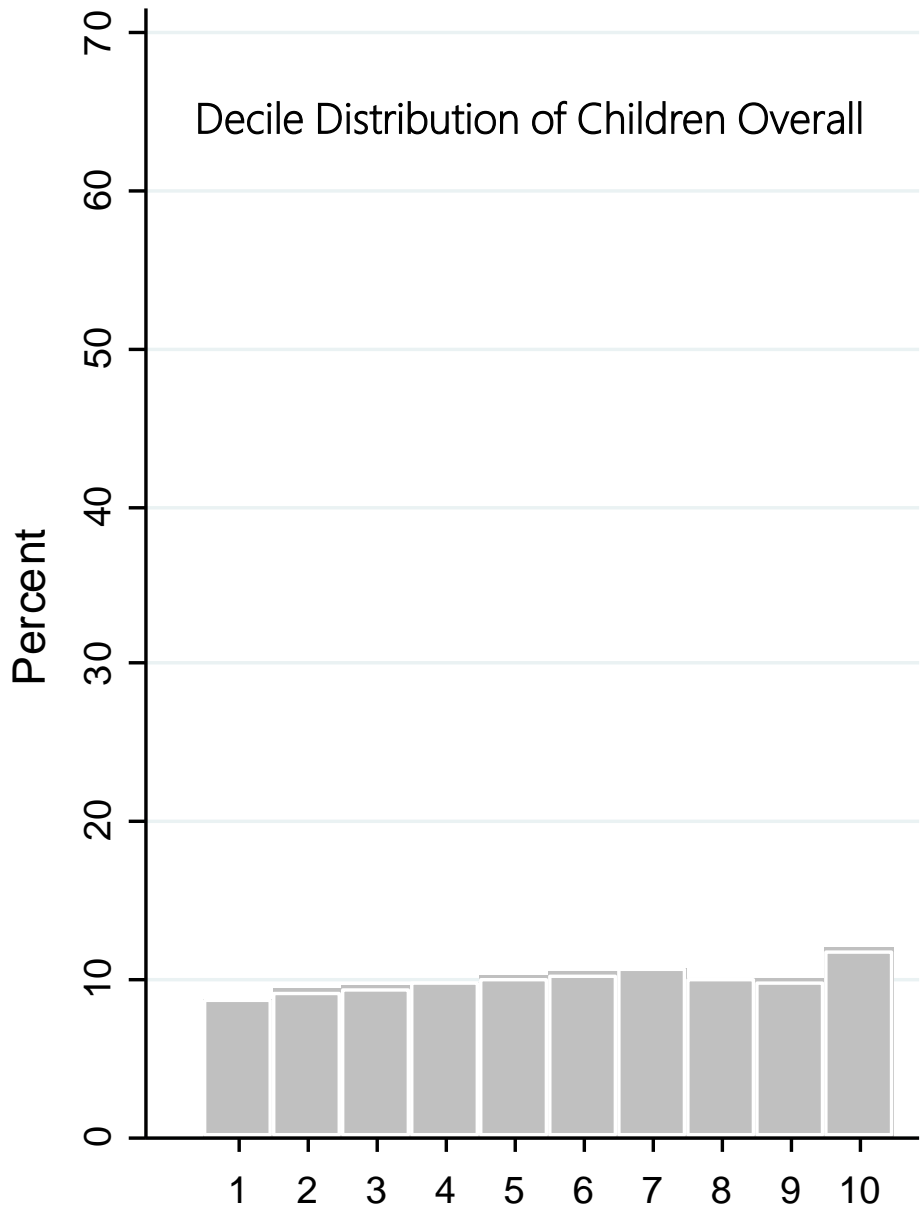
High risk white child/referrals (score of 10) are more likely to be screened out (21%) than high risk black and Hispanic child/referrals (similarly scored as 10s).

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

8 percentage point difference, highest to lowest

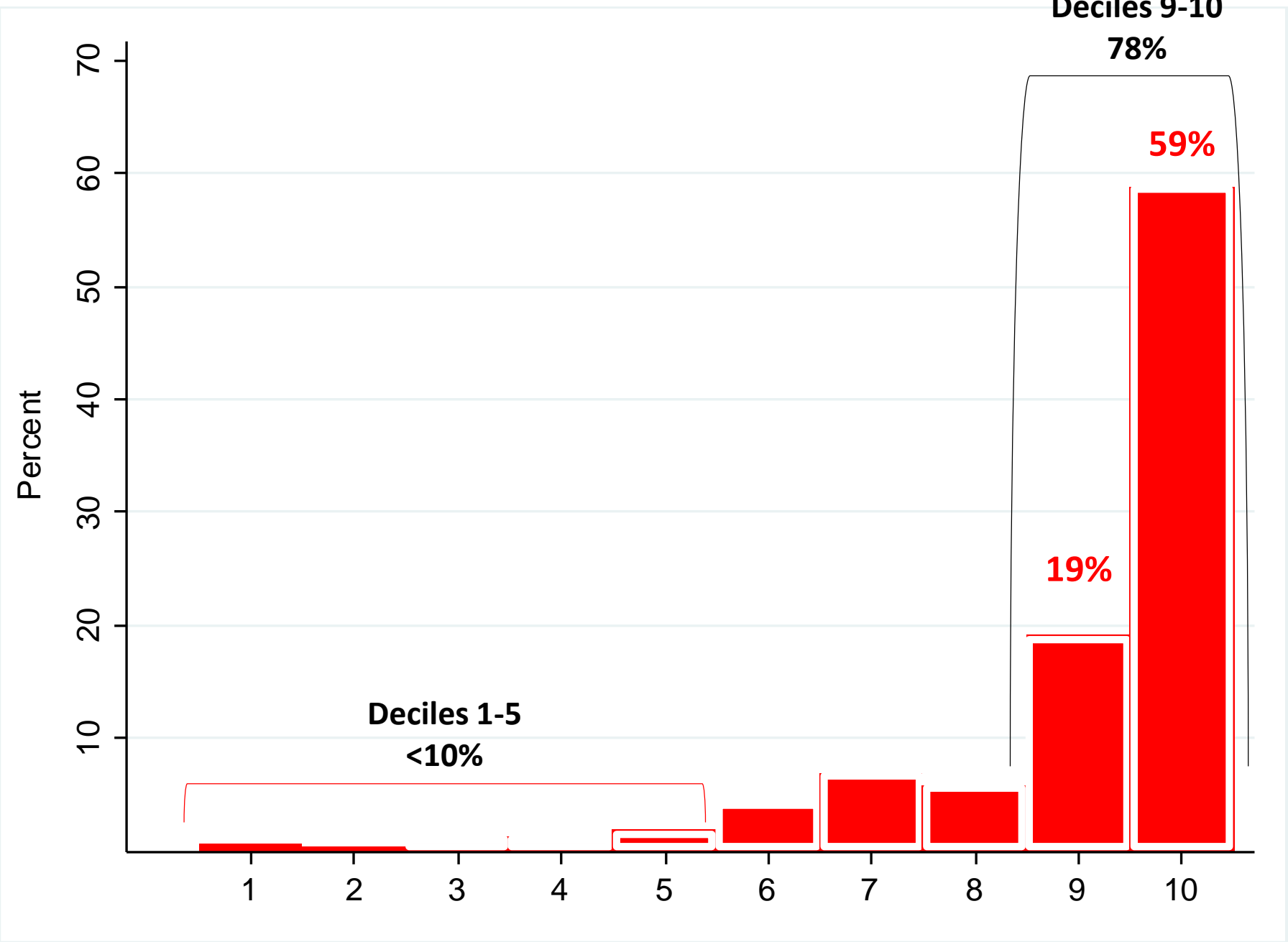
6 percentage point difference, highest to lowest

External Validation *[ongoing]*



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

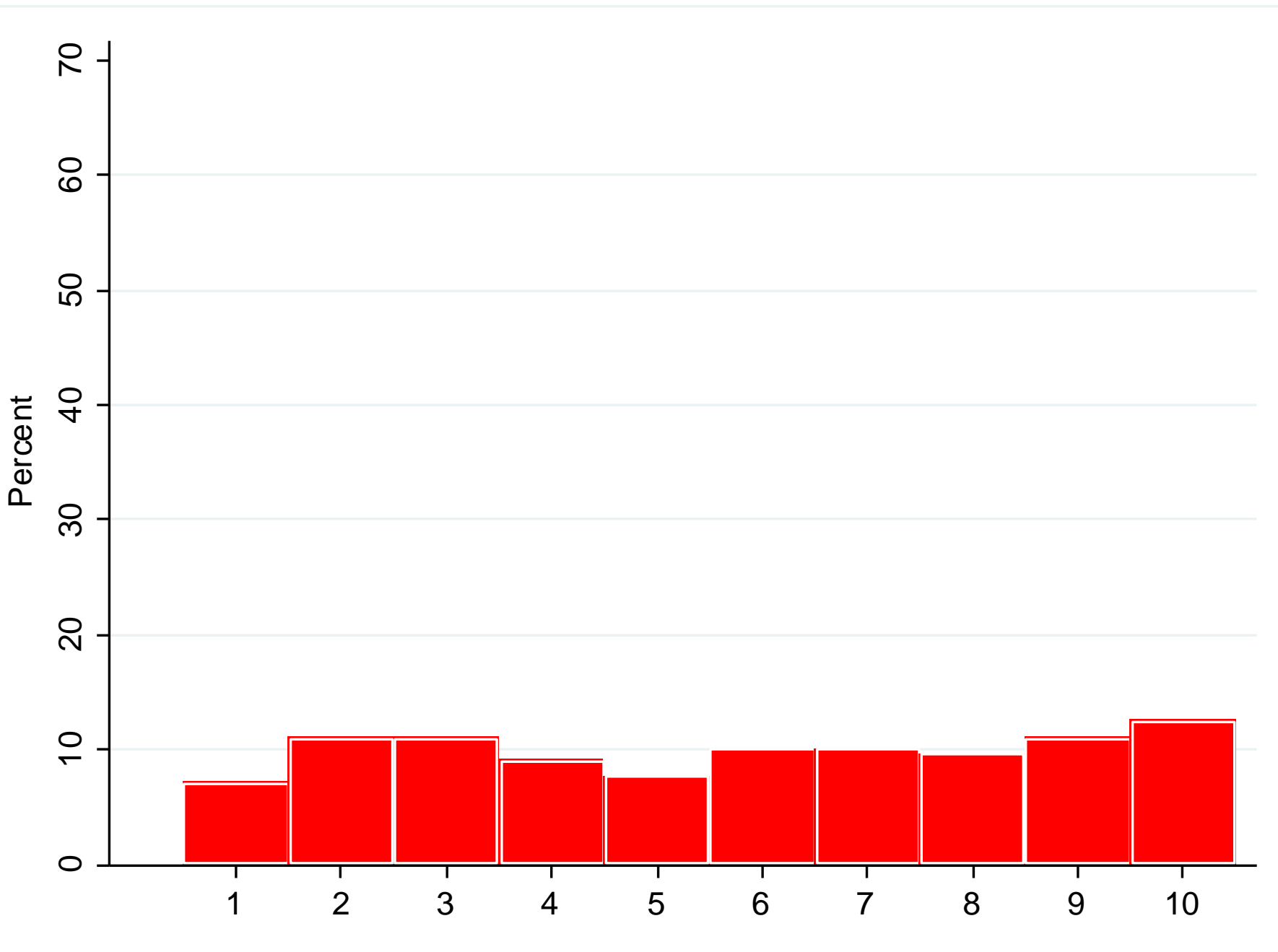
[maximum decile score]



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).

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



“...if there is a 50/50 chance that a newborn could get a communicable disease in the first 5 years of life based on known risk factors, public health professionals would jump at the chance of finding that newborn; they would not institute a generic public health preventative campaign at the community level in hopes that the newborn’s family might see that campaign. Public health professionals would use information on an individual newborn to customize a preventative program for that newborn and their family.”

L. Nguyen, American Journal of Public Health, 2014

Questions?

ehornste@usc.edu

