Prediction, Learning & Prevention in Child Welfare

Part 2

APHSA
American Public Human Services Association

INFLUENCE
BUILD
CONNECT

Thriving Communities Built on Human Potential

WWW.APHSA.ORG @APHSA1

1

Our Vision

Thriving Communities Built on Human Potential

Our Mission

American Public Human Services Association advances the well-being of all people by influencing modern approaches to sound policy, building the capacity of public agencies to enable healthy families and communities, and connecting leaders to accelerate learning and generate practical solutions together.

Because We Build Well-Being from the Ground Up

2

JUNE 7–10
POLICY AND PRACTICE
APHSA NATIONAL HEALTH AND HUMAN SERVICES SUMMIT
ARLINGTON, VA

AUGUST 3–6
PROGRAM INTEGRITY
NAPIPM EDUCATION CONFERENCE
PORTLAND, OR

AUGUST 23–26
SNAP/TANF
AASD/NASTA EDUCATION CONFERENCE
MILWAUKEE, WI

SEPTEMBER 13–16
TECHNOLOGY
ISM ANNUAL CONFERENCE
KISSIMMEE, FL

OCTOBER 4–7
ORGANIZATIONAL EFFECTIVENESS
NSDTA EDUCATION CONFERENCE
SPOKANE, WA

OCTOBER 18–21
LEGAL
AAHHSA EDUCATION CONFERENCE
COLUMBUS, OH

Thriving Communities Built on Human Potential WWW.APHSA.ORG @APHSA1

3

APHSA Analytics Committee

- https://aphsa.org/NC/NC/Guidance_and_Resources_Sub/data_sharing_analytics.aspx
  - Bitly: https://bit.ly/37Kvn95

1. *Analytic Capability Roadmap 1.0 for Human Service Agencies*, and a Reference Guide to Writing an Analytics RFP
2. *Roadmap to Capacity Building in Analytics*
3. *Guide to Data Management, Privacy & Confidentiality, and Predictive Analytics*
4. Still finalizing:
   - *Data Governance Value Proposition for Population Health*
   - 2-pager for Executives on Data Dictionaries

Thriving Communities Built on Human Potential WWW.APHSA.ORG @APHSA1

4

Part 1 Summary

- On January 14th, we heard from **New Allies**/Youth Villages & the **National Council of Crime & Delinquency** (NCCD)
- The **recording** & PPT is available at the bottom of the *Events* page on aphsa.org (or here: https://bit.ly/2RJUu5q)
- Summary
  - Preconditions of success
  - When to use decision support
  - How to develop & implement solutions
  - What it means to "get it right"

Thriving Communities Built on Human Potential

WWW.APHSA.ORG @APHSA1

5

Presenters

ERIN DALTON

Deputy Director

Office of Analytics, Technology and Planning, Allegheny County

CHRISTINA BECKER

Knowledge Mobilization Manager

APHSA

Thriving Communities Built on Human Potential

WWW.APHSA.ORG @APHSA1

6

Allegheny County Department of Human Services

Improving Decision Making in Child Welfare

7

7

What Key Decisions Are We Working On

Child welfare screening decision (Allegheny Family Screening Tool)

Homeless services (live this Spring)

Rethinking child abuse prevention (Hello Baby; live this Spring)

Elder abuse hotline

Mental health housing decisions

8

8

Process Non-Negotiables

- Commitment to Implement
- Do Something that Matters
- Competitive Procurement
- Built in the Public Domain (we own the model etc.)
- Ethical Review
- Model Fairness & Discrimination Review
- Validation
- Stakeholder Input
- Community Engagement
- Willingness to Modify
- Evaluation
- Commitment to Improve
- Transparency

It was 2014….

Sick over what looked like bad decision making in call screening (it's more complicated that)

Having no decision making support at this important stage

Received 16 proposals (https://www.alleghenycounty.us/Human-Services/Resources/Doing-Business/Solicitations/Awarded-Contracts/Award-Details---023.aspx)

The field of data ethics and even data science was just emerging

We could keep status quo, we could implement Structured Decision Making, we could see if a predictive risk modeling approach might work better

Homelessness Rabbit Hole

What I think we thought back then and what we know now

11

Improving Response to Homelessness

30,000 calls

Today, we use an assessment

What if we use the data we already have

12

12

Recent Evaluation of the Assessment Tool

Test Retest Reliability
fell below the accepted cutoff for good reliability

Inter-Rater Reliability
was inadequate on 4 items

Construct Validity
Several items were not strongly associated or were associated in an unexpected direction

Predictive Validity
Only marginally associated with the likelihood of re-entering homeless services

13

Model Predictors & Outcomes

Demographics (Age & Gender)	Homeless	Child Welfare
Jail	Courts	Probation
Juvenile Probation	Assisted Housing	Behavioral Health

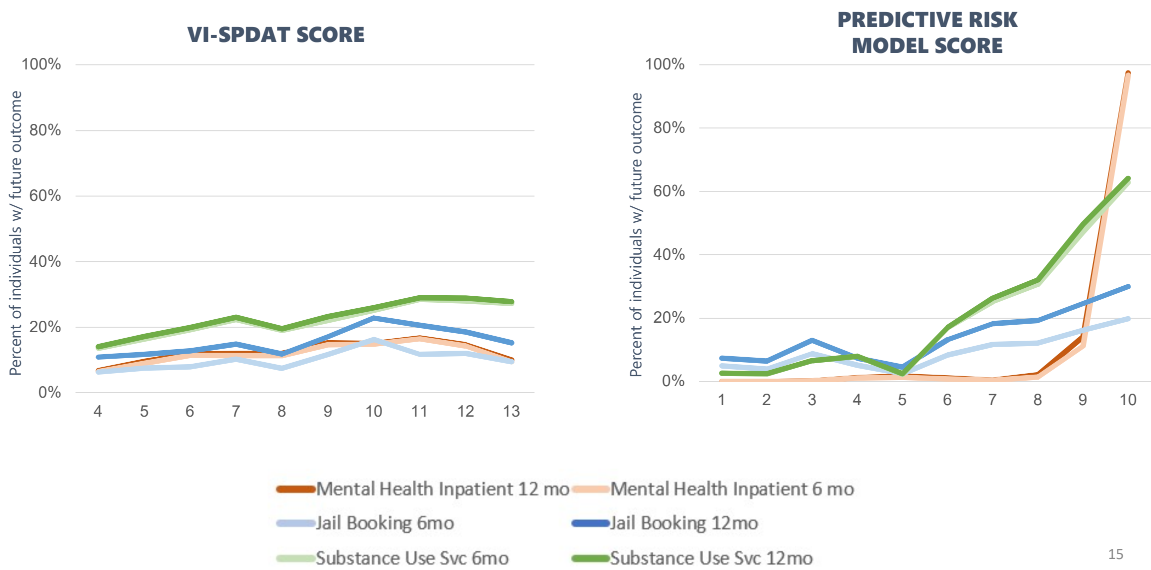
Outcomes

Likelihood of:
- **4 or more Emergency Department visits**
- **Mental Health Inpatient stay**
- **Jail Booking**
- **Substance Use Treatment**

14

14

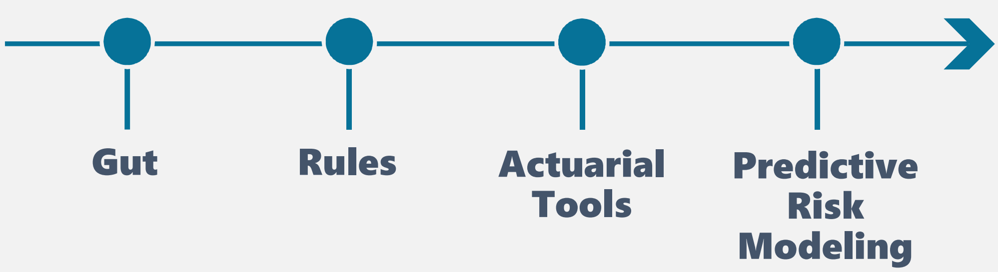
Current Assessment Vs Predictive Risk Model



15

15

Improving Decision Making: Typical Progression



16

16

Back to Child Welfare

We decided the *try* a predictive risk model

17

The Team

Research Team
- Rhema Vaithianathan, Auckland University of Technology
- Emily Putnam-Hornstein, USC
- Alexandra Chouldeshova, CMU

Ethical Review
- Tim Dare, University of Auckland
- Eileen Gambrill, UC Berkeley

Validation Study
- Rachel Berger, Chief, Child Advocacy Center Children's Hospital of Pittsburgh

Evaluators

Process
- Hornby-Zellar Associates

Impact
- Stanford University

Technology Implementation
- Deloitte (initially)
- Data Warehouse Consultants (now)

18

18

A report of child abuse is made every 10 seconds in the US, involving 6.6 million children per year

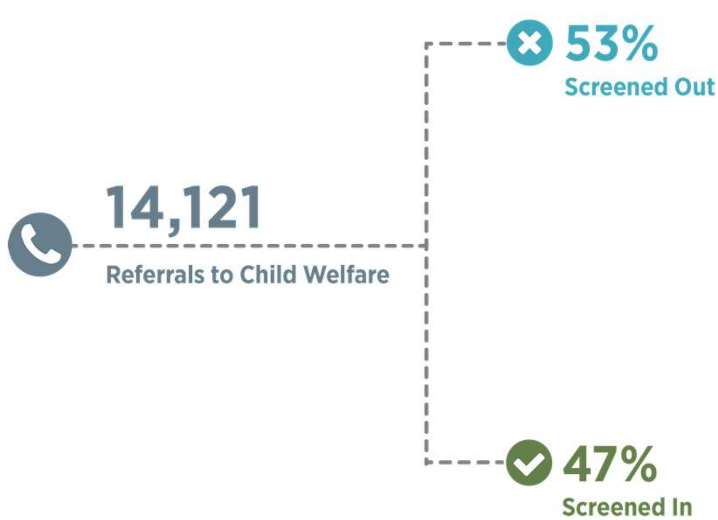
37% of children in the US will experience a child abuse investigation at some point in their childhood

We are not the police. We don't have resources to respond to every report

Consequences are tremendous

19

Improving Hotline Decision-Making



20

Researchers built a screening model based on information that we already collect

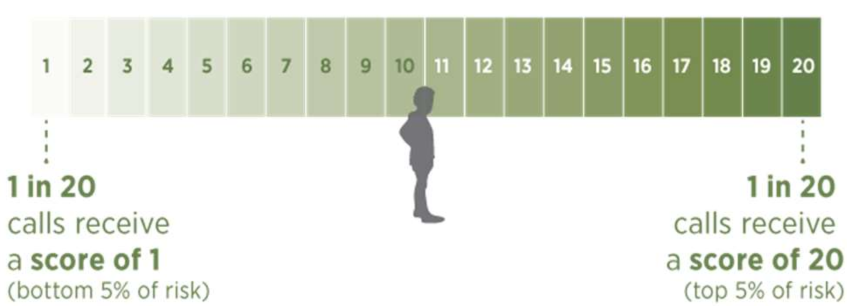
They identified more than 100 factors that predict future referral or placement

To test if the model might improve the accuracy of screening decisions, we scored thousands of historical maltreatment calls and then followed the children in subsequent referrals to see how often the model was correct…

21

Developing a Screening Score

- **The screening score is from 1 to 20**
- **The higher the score, the higher the chance of the future event** (e.g., abuse, placement) according to the data



22

The Results: Out-of-Home Placements

23

The Results: Out-of-Home Placements

24

25

External Validation

Relative risk of a child whose placement score at referral is 20 vs. a child who scores 1 is…

- **21 times** more likely to attend for a self-inflicted injury
- **17 times** more likely to attend for physical assault
- **1.4 times** more likely to attend for an accidental fall

26

What would you do?

Would you try to implement?

27

Implementation

- Live since August 2016
- Fixed bugs in November, 2016
- Major changes to model, business processes & policies, November, 2018

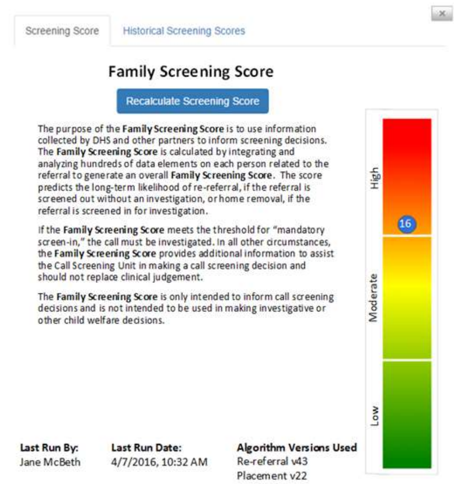
So far:

- Viewed in 100% of cases
- No increase in investigations but an increase in cases opened
- Not replacing clinical judgement but getting more concurrence
- Low risk protocol now in place
- Looking at other decisions

28

Significant Changes in Version 2

- What we are predicting
- Underlying data used
- Modeling method
- Added a low risk protocol
- Immediate feedback from the workforce
- Enhanced quality assurance



29

Modeling Comparisons

- This table shows the **positive predictive value (PPV)** and **true positive rate (TPR)** for the four models with respect to the high- and low-risk protocols on the test data.
- All models flagged ~25% of test referrals as high-risk.
- For the LASSO model, 47.6% of referrals flagged as high-risk experienced a removal – and 53.7% of referrals experiencing a removal would have been flagged as high-risk.

	High-Risk Flag				Low-Risk Flag			
	Logistic Regression	LASSO (AFST V2)	Random Forest	XG Boost	Logistic Regression	LASSO (AFST V2)	Random Forest	XG Boost
Proportion of referrals that receive the flag	23.4%	23.8%	25.1%	26.2%	9.2%	4.1%	2.8%	2.9%
Proportion of referrals flagged where child ends up placed within 2 years (PPV)	35.4%	47.6%	47.6%	46.2%	16.4%	7.6%	5.9%	4.4%
Proportion of all referrals where child ends up being placed, who are flagged	39.3%	53.7%	56.6%	57.4%	7.2%	1.5%	0.8%	0.6%

30

Screening and Acceptance Rates

- At the screening stage, referrals are **more likely to be screened-in when the AFST score is higher**.
- This is especially true for the High-Risk and Low-Risk Protocols, which require an explicit supervisor override to screen out (if High) or in (if Low).
- While investigators do not receive or know the AFST scores, **investigations with higher scores also tend are more likely to end up being accepted** for case openings.

LASSO model (2019)	Call Screening			Investigation	
AFST Score Tier	Count Screened	Screened-In for Investigation	Screen-In Pct.	Accepted for Services	Acceptance Pct. (among investigations)
High-Risk Protocol	1265	943	75%	263	28%
High Range (15-20) (No Protocol)	1854	915	49%	299	33%
Medium Range (10-14) (No Protocol)	2405	1003	42%	276	28%
Low Range (1-9) (No Protocol)	1641	496	30%	104	21%
Low-Risk Protocol	193	36	19%	5	14%
No Score Generated	949	176	19%	38	22%
All GPS Referrals	**8307**	**3569**	**43%**	**985**	**28%**

Notes: Includes all incoming GPS-only referrals between 2/1/19 - 10/26/19, with truancy-only court notices omitted.

31

Impact Evaluation

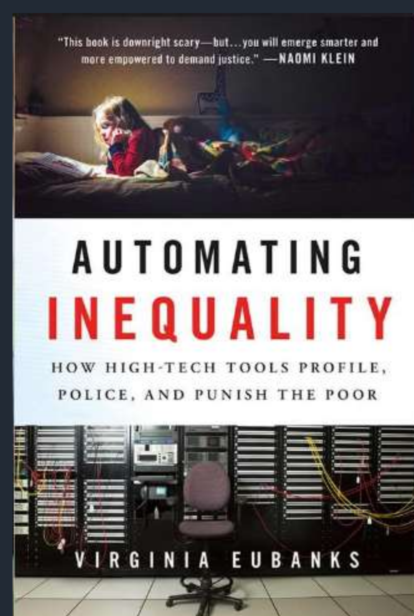
"Implementation of the AFST saw no adverse consequences and increased the accurate identification of children who needed further intervention services, without increasing the workload on investigators."

32

Impact Evaluation

- Increased the identification of children determined to be in in need of further child welfare intervention.
- Led to reductions in disparities of case opening rates between black and white children.
- Did not lead to increases in the number of children screened-in for investigation.
- No evidence that the AFST resulted in greater screening consistency.

33



34

Some Lessons Learned

- Modeling is hard but it's not the hardest part
- The policies, procedures around the model are critical
- You don't need super-rich integrated data need to do this
- You can probably do better than the way you are making decisions now
- There is not one (or any agreed upon) definition of fairness, we look at multiple approaches & their trade-offs
- We don't make the hard decisions alone

35

alleghenycountyanalytics.us

Erin.Dalton@alleghenycounty.us

36

Q & A

Thank you!

37