



**CENTRE FOR
SOCIAL DATA ANALYTICS**

Predictive Analytics in Child Welfare

Rhema Vaithianathan

Three Approached to identifying families in need or at risk

- Clinical Judgement
- Threshold Models (structured decision models)
- Predictive Risk Modelling

Threshold or Structured Models

Identify people on the basis of a

list of characteristics

observed by a frontline professional

Family Start Referral Criteria

List A sets out the main criteria for Family Start. Families need to experience challenges in one of these areas to be referred.

List B sets out other areas where families may have challenges. If you feel that a parent has high needs but does not 'fit' into the Family Start criteria then explain, using indicators in List B, why an exception should be made for this family.

LIST A. Family Start criteria	DESCRIPTION
Mental health issues	<p>Either parent/carer has a mental health problem, for example:</p> <ul style="list-style-type: none">• Post-natal depression• Anxiety• Depression• Self-harm or suicidal tendencies• Other (specify) <p>Is this current or historic?</p> <p>Are they receiving any assistance or professional services for this?</p>
Difficulties with drugs, alcohol or gambling	<p>Either parent/carer has a problem with one of the following which affects parenting ability:</p> <ul style="list-style-type: none">• Alcohol use• Illicit drug use• Gambling <p>Are they are receiving any assistance from other professional services for this?</p>
Childhood history of abuse	Either parent/carer experienced abuse as a child or young person.

Admissions criteria for a NZ home visiting program called "Family Start"

What's wrong with threshold models?

- Operators distort inputs to get the outcomes they want
- Expensive to administer – so **only a small population can be screened**

Predictive Risk Modelling (PRM)

Automatic risk scoring tool which generates a risk score for an adverse event based on large administrative dataset

Advantages of PRM

- Cost effective screening of large populations e.g. all hospital admissions or all births
- No human factor involvement
- Risk score is a “continuum”
- Built from the same data as the population

Disadvantages

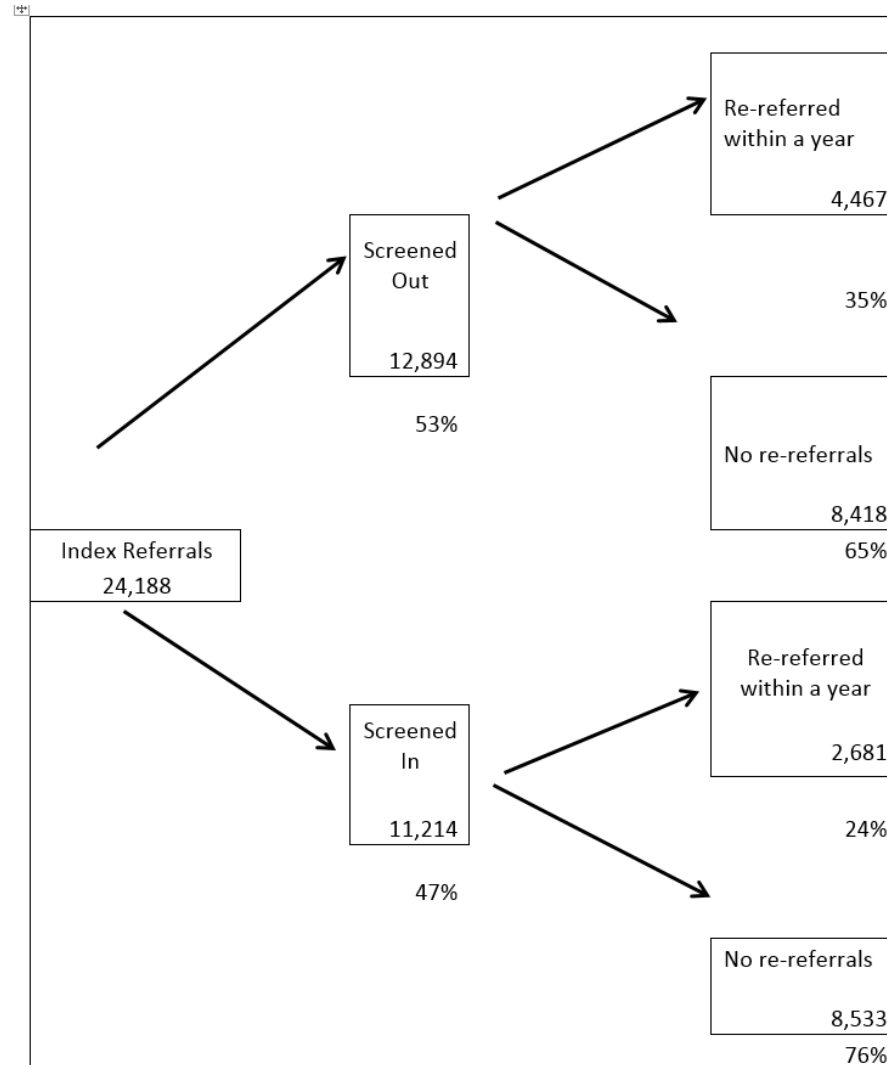
- “Black Box”
- Resistance by clinicians and frontline staff
- Only as good as the preventive intervention
- Risk is not the same as “amenability” to the intervention

How could a predictive model work?

Allegheny Call Screening Model
(Preliminary Proposal)

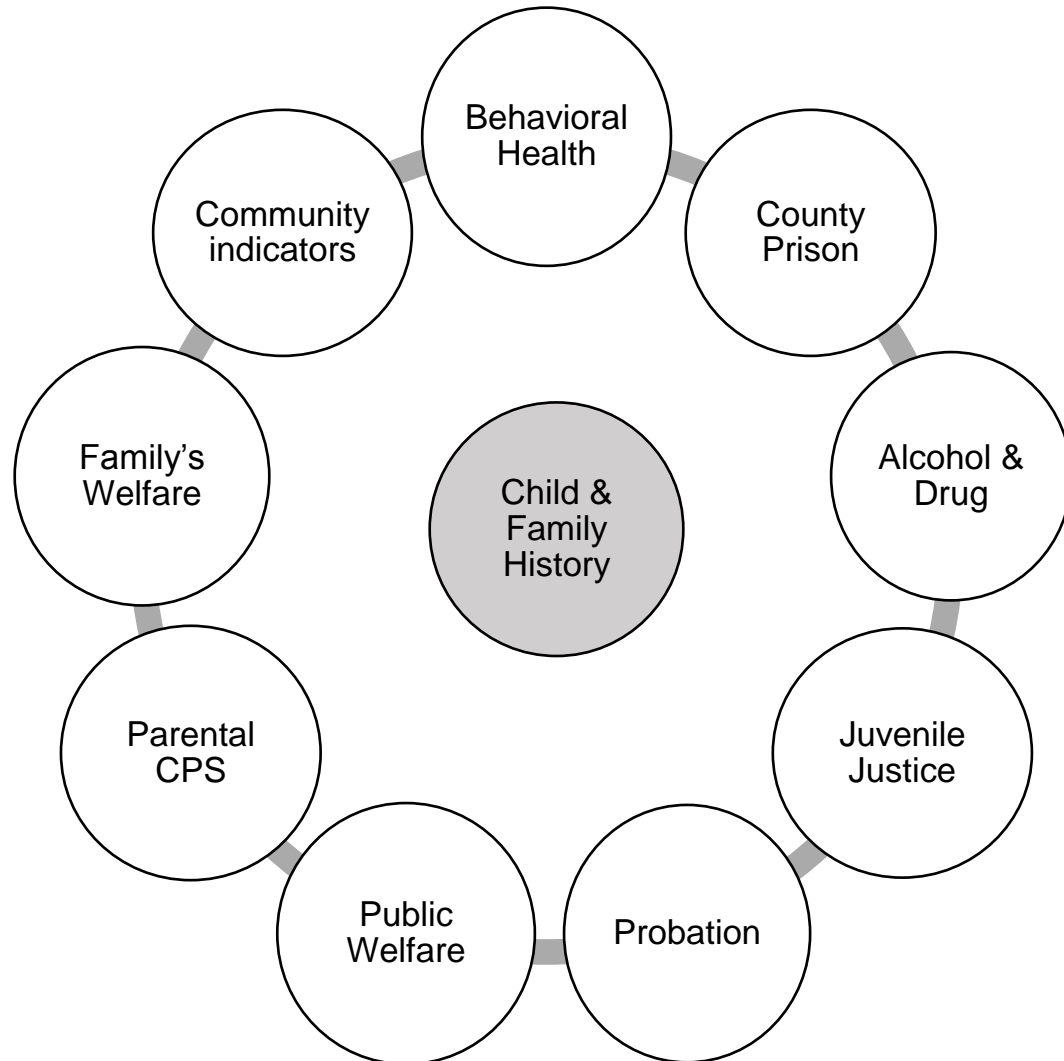


Current practice when an allegation of abuse or neglect is made to the hotline



Data used is for all children named as a victim in an index referral between March 2010 and Feb 2013

How a predictive model works



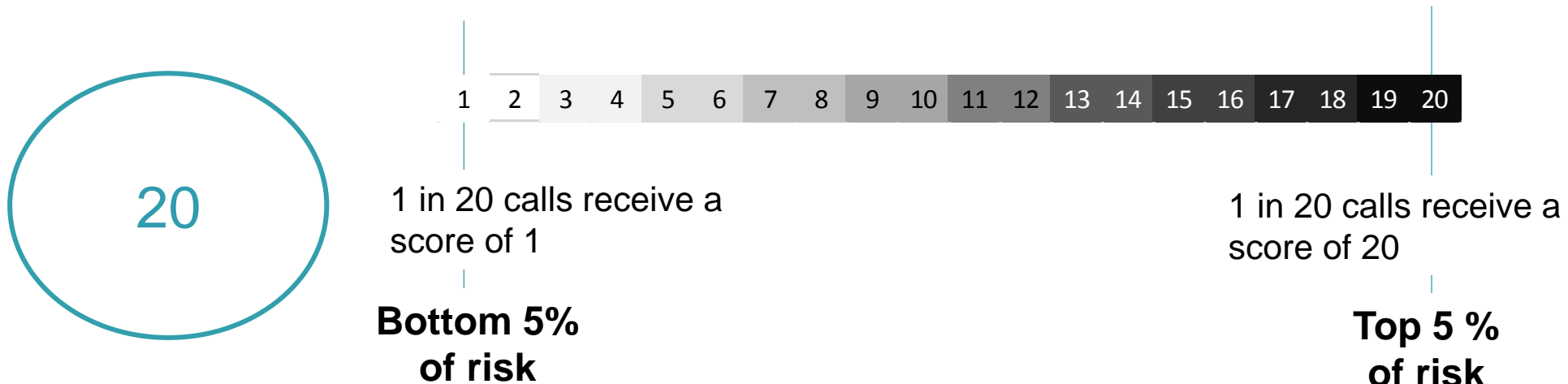
The system harvests data from the data warehouse...

A Risk Score is produced for each child named in the call*

*Only if an MCI_ID is successfully established. The referral data used to build the model covers the period from 25 Aug 2008 to 13 March 2015, and there are 58,801 referrals (calls) and 50,076 (unique) victims in total. 2,236 victims (4.5%) did not have established MCI_ID

The Allegheny Family Screening Score

- The predicted risk score is from 1 to 20
- It tells us how likely the child is to **have a placement in the 2 years if screened in or re-referral if screened out** following the maltreatment call
- The higher the score, the higher the chance of placement



Operationalising the Score

Auto Screen In

= 1 in 5 referrals

High Risk

= 1 in 4 referrals

Medium Risk

= 1 in 4 referrals

Low Risk

= 1 in 4 referrals

SHOWN TO SCREENING STAFF ONLY

- Only auto-screen in have to be screened in
- Screening staff are encouraged to continue to use their judgement
- Training will emphasise that this is only a supplementary tool

We undertook scoring of past referrals and followed them to see how well they did...

Auto Screen In

vs

Low Risk


- ✓ 7 times as likely to have ever been involved in a critical event (ACT 33)
- ✓ 18 times as likely to be in placement within a year
- ✓ 3 times as likely to be re-referred within 2 years if screened out
- ✓ Current practice? **ONLY** 3 times more likely to be screened in

Allegheny Family Screening Score

- Given to call screening staff ONLY
- Independent ethical report provided to Allegheny by Tim Dare and Eileen Gambrill
- Process evaluation undertaken by HZA
- Impact evaluation to be undertaken by Stanford
- Evaluations completed end of 2017

What are the steps for Colorado to establish an analytic tool?

1. Decide some options about when you want the tool to be used (at call screening, intake, service initiation) and who will act on it.
2. Review what data is available at the time you want to deploy an analytic tool. Even if only welfare data available, might be sufficiently powerful.
3. Extract a long (5 year) retrospective data from your systems and structure the data.

- 
4. Decide what you will predict – indicator of “need” or something you actively want to prevent.
 5. Build the algorithm and test for sensitivity, specificity, area under ROC. Compare with current practice. Test against ethical criteria. Test impact on disparities.
 6. Work with IT providers to create visualisation tools and implement the algorithm.
 7. Engage ongoing independent evaluation and monitoring.

Elements to Commissioning Analytics

1. Analytics is only a small part of the challenge.
2. When to use the tool? Who should see the score?
3. How to judge if sufficiently accurate ?
4. How to train staff to overcome bias.
5. Need independent ethical report
6. Need to consider impact on disparities



Visit us at the Centre for Social Data Analytics

www.csda.aut.ac.nz

Questions?

