

# Implementing a Data Visualization Tool for the Frontline in Allegheny County Pennsylvania

Nan Jiang



# Research Team

## ***LEAD***

Rhema Vaithianathan, Auckland University of Technology

## ***Social Work***

Emily Putnam-Hornstein, USC

Irene de Haan, University of Auckland.

Eileen Gambrill, UC Berkeley

## ***Economists***

Marianne Bitler, UC Irvine

Tim Maloney, Auckland University of Technology

Nan Jiang, Auckland University of Technology

## ***Ethics***

Tim Dare, University of Auckland



# Reactive Predictive Risk Model

- What is a ‘Reactive’ PRM?
- What is being predicted?
- Research Dataset
- Model Development
- Model Performance
- Implementation Process

# Reactive Predictive Risk Models (PRMs)

- Predictive Risk Analytics: Integrate and utilize data **already collected** (i.e. could be accessed manually, in principle) about an individual to evaluate the possibility of future adverse outcomes
- **Reactive** vs. **Proactive** use of the tool. A reactive use is when the tool is used *subsequent* to a referral to child welfare – i.e. only when we have *prima facie* grounds to think a child may be at risk. This avoids some of the *consent issues* which would arise if the tool were used at birth in a proactive manner.
- Rights of those who would normally be given an opportunity to consent must be balanced against rights of vulnerable children.

# What is being predicted?

ALL the Referral calls from KIDS (Key Information Demographic System) were pulled and the data includes:

- 1) Who are the victims, the perpetrators, the parents and all other people identified to be associated with this referral;
- 2) Their respective age, gender, ethnicity profile, and address;
- 3) Where did that call come from, reported by a bystander, school, from hospital, by daycare etc.;
- 4) The initial call screening decision, service decision and the final finding of the case

# A concrete example....

A referral comes in....



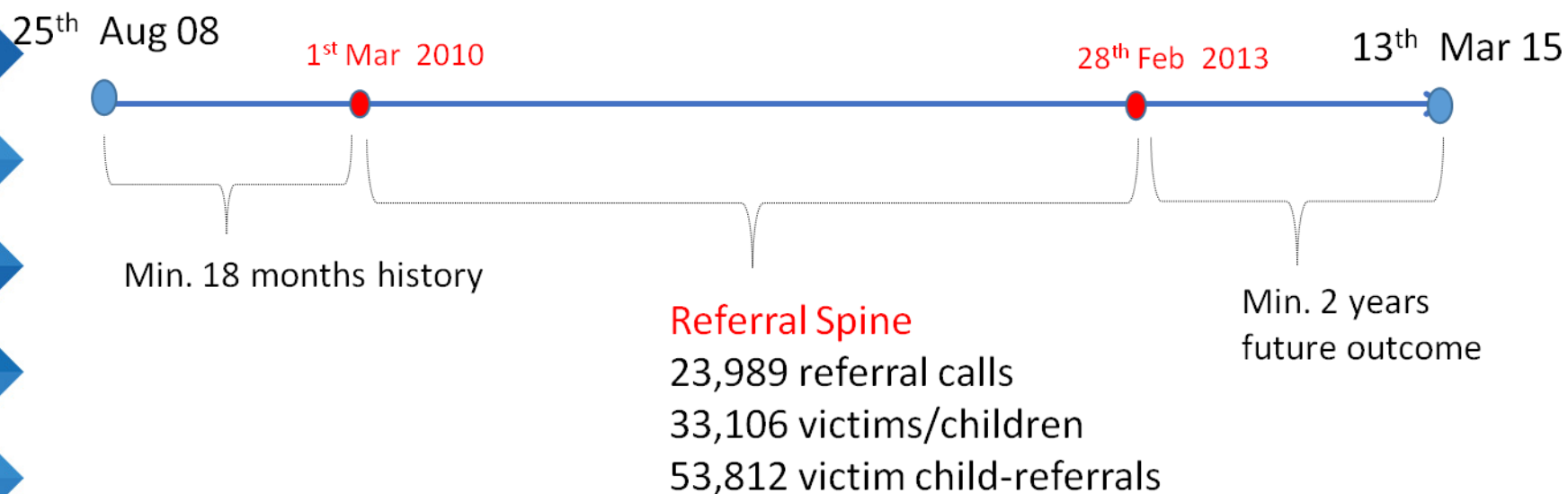
KIDS started

25<sup>th</sup> Aug 08

13<sup>th</sup> Mar 15

58,801 referrals (calls)  
155,136 (unique) person  
67,612 (unique) victims children

# What is being predicted?



\*4,408 victims children (6.5%) did not have established MCI\_ID.



**Referral  
Calls**  
23,989

## Screening

**Screen Out**  
12,461 (52%)

**Screen In**  
11,528 (48%)

## Service Decision

**Service**  
5,957 (52%)

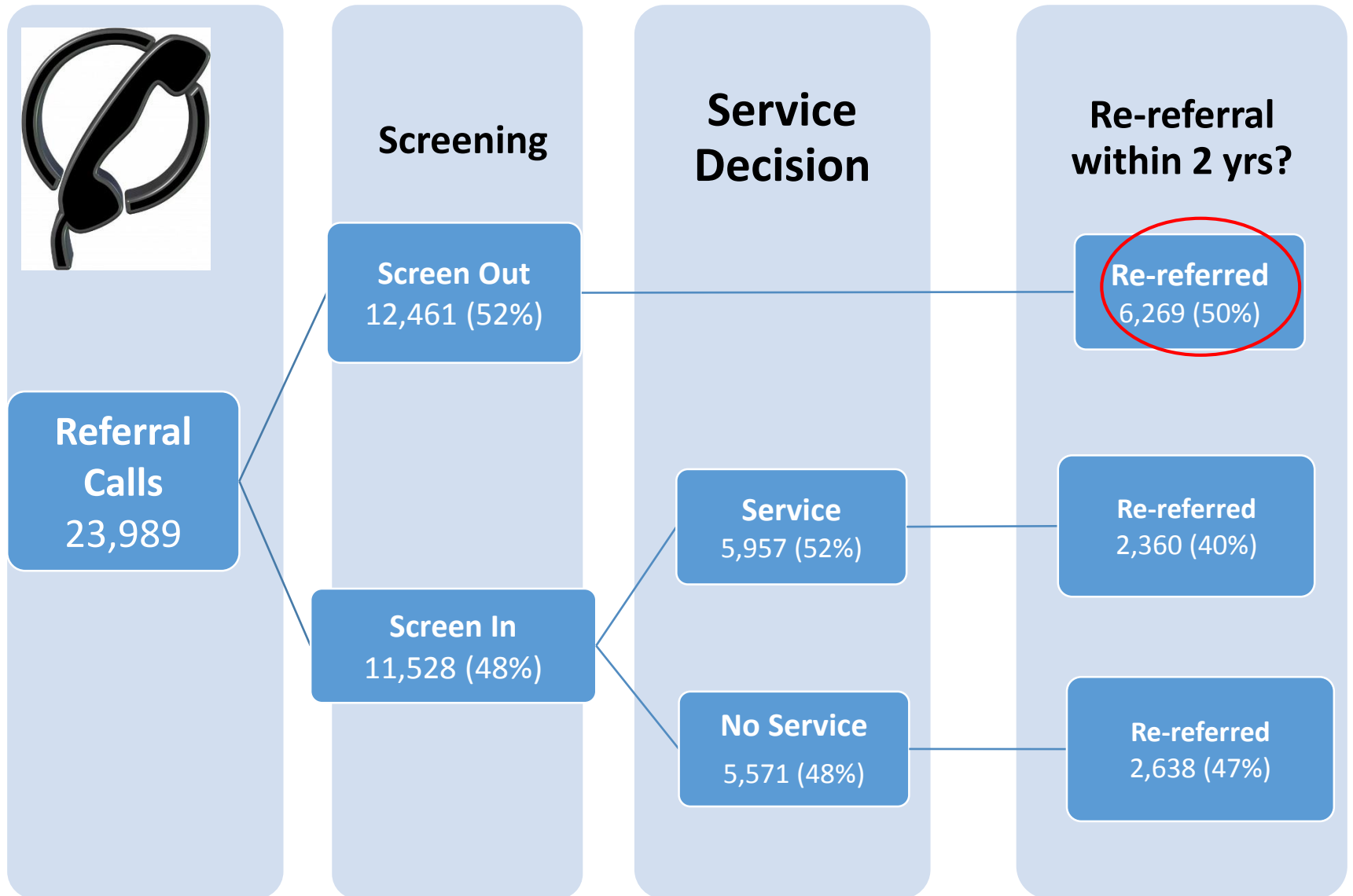
**No Service**  
5,571 (48%)

## Re-referral within 2 yrs?

**Re-referred**  
6,269 (50%)

**Re-referred**  
2,360 (40%)

**Re-referred**  
2,638 (47%)







**Referral  
Calls**  
23,989

## Screening

**Screen Out**  
12,461 (52%)

**Screen In**  
11,528 (48%)

## Service Decision

**Service**  
5,957 (52%)

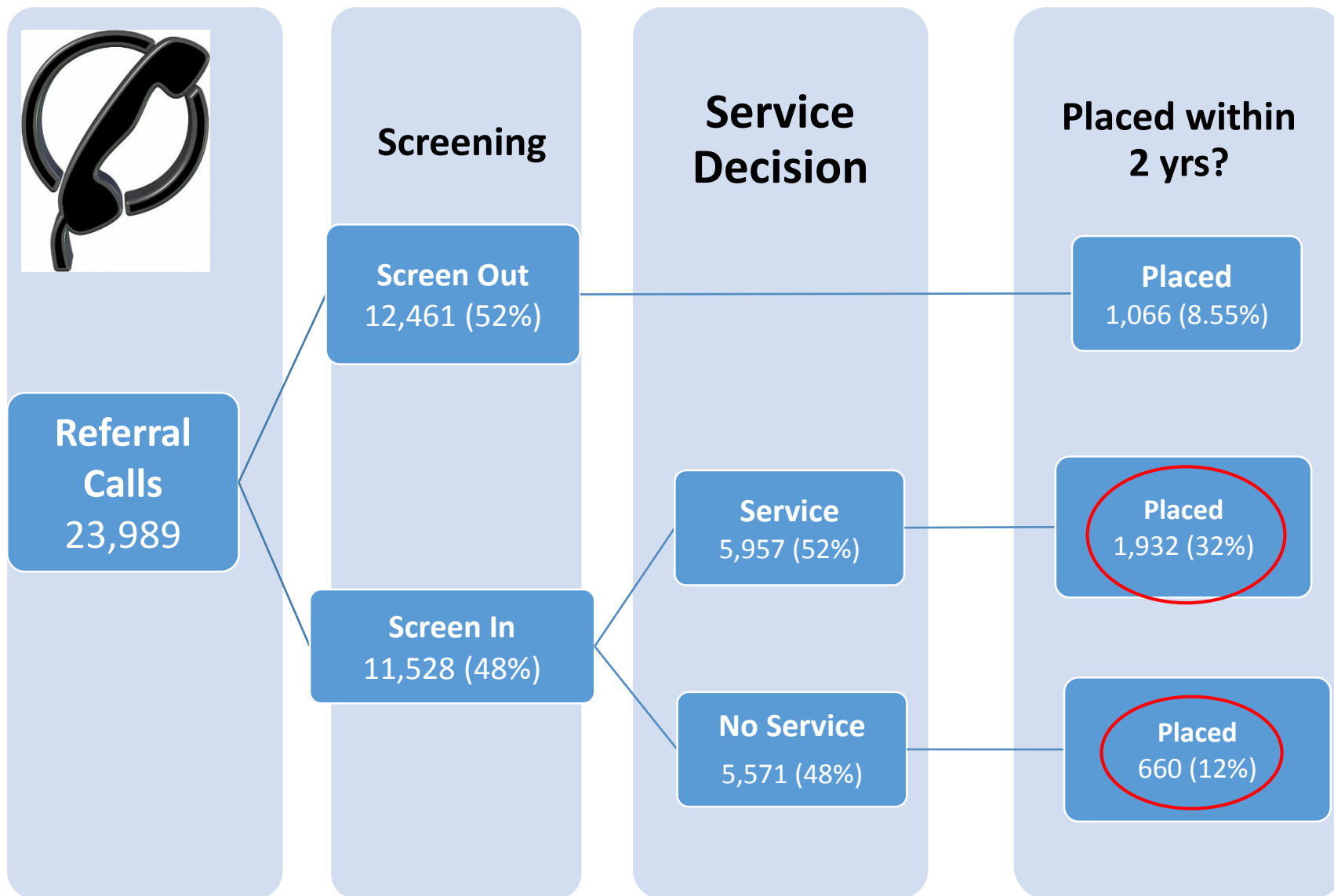
**No Service**  
5,571 (48%)

## Placed within 2 yrs?

**Placed**  
1,066 (8.55%)

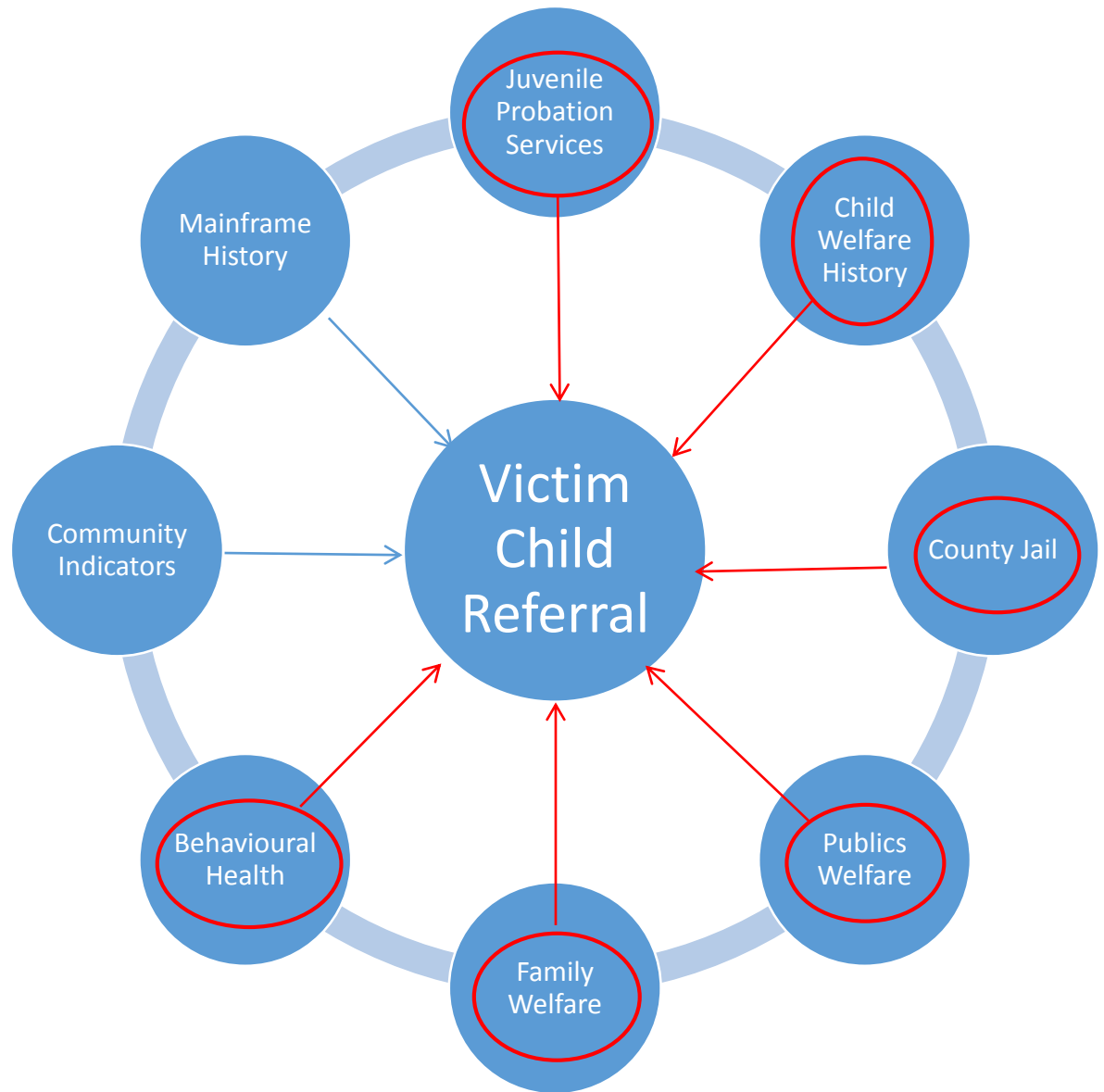
**Placed**  
1,932 (32%)

**Placed**  
660 (12%)



# Dataset Construction





\*Only if an MCI\_ID is successfully established



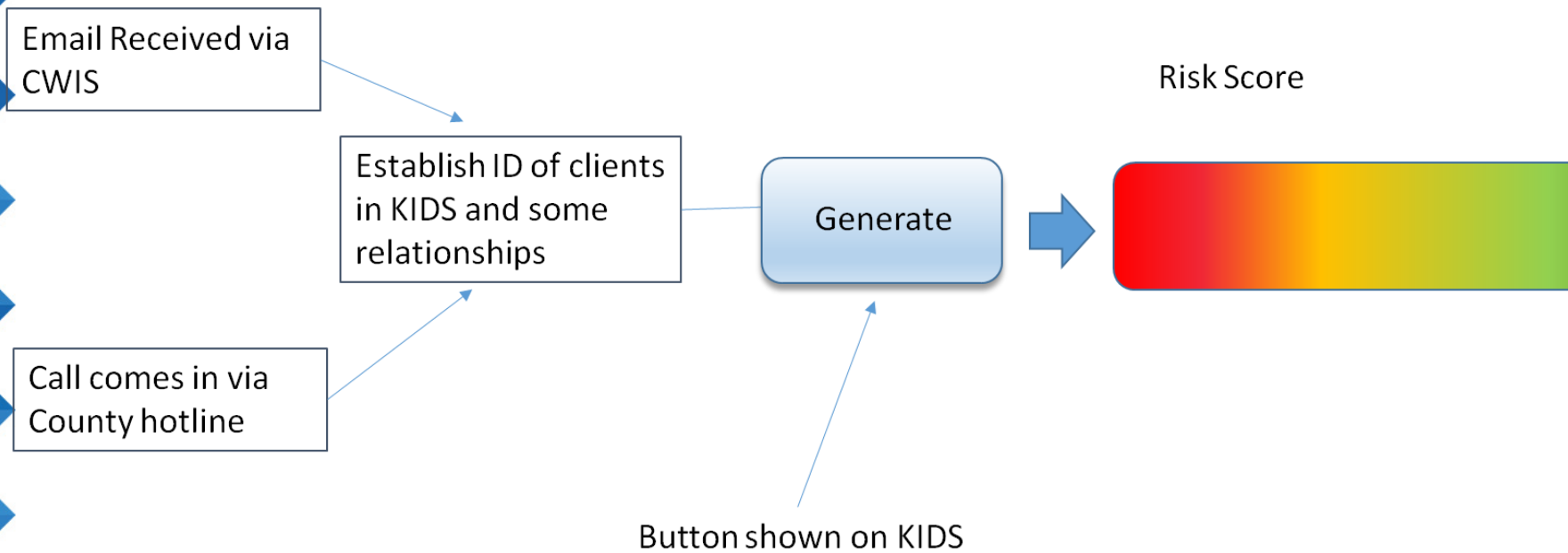
# Model Development

- We built initial models by selecting blocks of variables and testing their statistical significance
- Then each of these blocks were interacted with age of the victim, race, source of referral
- The final model was then built on a random 70% of the sample and tested on the other 30%

# Model Performance

20 quantiles of risk score	Re-referred	Placed
1	10.63%	0.86%
2	23.38%	2.50%
3	33.96%	3.64%
4	27.41%	3.88%
5	45.45%	3.74%
6	32.52%	5.65%
7	52.29%	5.56%
8	51.47%	6.03%
9	58.06%	5.61%
10	56.41%	4.96%
11	68.87%	8.04%
12	59.62%	10.26%
13	69.23%	10.57%
14	63.49%	8.25%
15	64.15%	14.29%
16	79.46%	16.79%
17	72.90%	19.85%
18	83.93%	23.63%
19	87.18%	30.77%
20	92.00%	39.89%
ROC Area	75%	76%
Testing Sample N	8,110	8,245

# Implementation



# Service Elements...

What  
should  
happen?

How should the  
score be seen by  
the frontline  
social worker?



What should and  
**should not** happen  
when call comes in?



*Depends on accuracy  
of the model vis-a-viz  
current practice.*





Deploy **after**  
screening decision.  
High risk children  
who are screened  
out offered voluntary  
services

Allow  
supervisors  
only to see  
the score.

Provide to  
supervisors and  
caseworkers before  
any decisions are  
made.

CONFIDENCE IN RISK SCORING TOOL (compared to usual practice)

Moderate  
Confidence

High  
confidence



# Data Visualization Tool

What should the case worker see at the same time as they are provided with the risk score?



# Implementation

Risk Scores – possible alternative presentations



1. Showing precise location of the score and possibly number



2. Showing classification



3. Words and colours

Client ID: 12354

Client ID: 6789

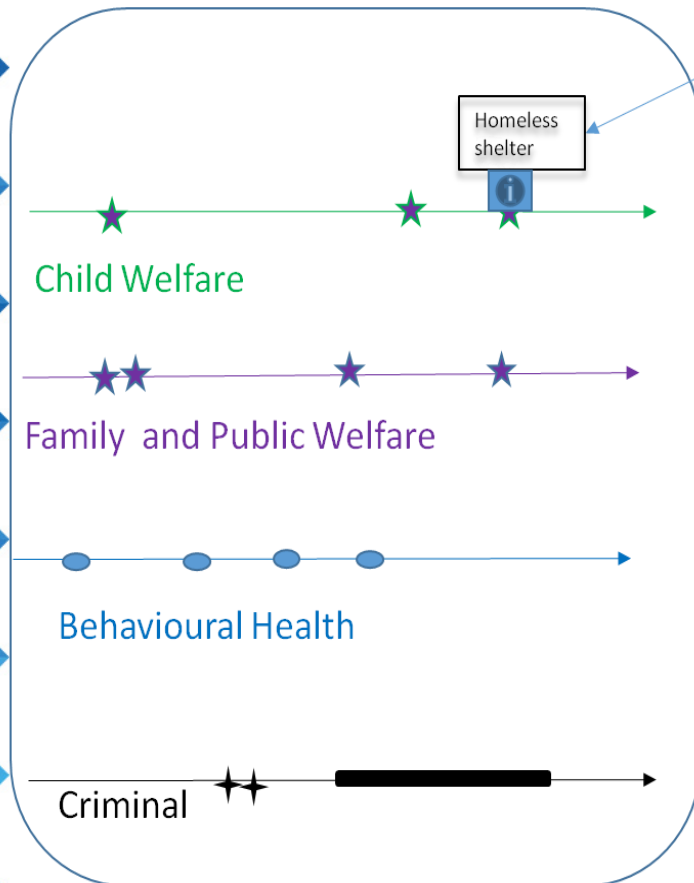
Client ID: 1011



Can view for individual clients in a referral or overlay for all clients

Hover buttons will allow more information on the event

Can save this picture to attach to reports



Save