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
25th Aug 08

13th Mar 15

58,801 referrals (calls)
155,136 (unique) person
67,612 (unique) victims children
What is being predicted?

25th Aug 08

1st Mar 2010

Min. 18 months history

28th Feb 2013

13th Mar 15

Referral Spine
23,989 referral calls
33,106 victims/children
53,812 victim child-referrals

*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%)

Referred 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

- Demographics
- Source
- Referral Summary
- Response
- Referral history
- Placement history
- Outcomes within 2 years

Victim Child Referral
*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

<table>
<thead>
<tr>
<th>20 quantiles of risk score</th>
<th>Re-referred</th>
<th>Placed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.63%</td>
<td>0.86%</td>
</tr>
<tr>
<td>2</td>
<td>23.38%</td>
<td>2.50%</td>
</tr>
<tr>
<td>3</td>
<td>33.96%</td>
<td>3.64%</td>
</tr>
<tr>
<td>4</td>
<td>27.41%</td>
<td>3.88%</td>
</tr>
<tr>
<td>5</td>
<td>45.45%</td>
<td>3.74%</td>
</tr>
<tr>
<td>6</td>
<td>32.52%</td>
<td>5.65%</td>
</tr>
<tr>
<td>7</td>
<td>52.29%</td>
<td>5.56%</td>
</tr>
<tr>
<td>8</td>
<td>51.47%</td>
<td>6.03%</td>
</tr>
<tr>
<td>9</td>
<td>58.06%</td>
<td>5.61%</td>
</tr>
<tr>
<td>10</td>
<td>56.41%</td>
<td>4.96%</td>
</tr>
<tr>
<td>11</td>
<td>68.87%</td>
<td>8.04%</td>
</tr>
<tr>
<td>12</td>
<td>59.62%</td>
<td>10.26%</td>
</tr>
<tr>
<td>13</td>
<td>69.23%</td>
<td>10.57%</td>
</tr>
<tr>
<td>14</td>
<td>63.49%</td>
<td>8.25%</td>
</tr>
<tr>
<td>15</td>
<td>64.15%</td>
<td>14.29%</td>
</tr>
<tr>
<td>16</td>
<td>79.46%</td>
<td>16.79%</td>
</tr>
<tr>
<td>17</td>
<td>72.90%</td>
<td>19.85%</td>
</tr>
<tr>
<td>18</td>
<td>83.93%</td>
<td>23.63%</td>
</tr>
<tr>
<td>19</td>
<td>87.18%</td>
<td>30.77%</td>
</tr>
<tr>
<td>20</td>
<td>92.00%</td>
<td>39.89%</td>
</tr>
</tbody>
</table>

ROC Area: 75% 76%

Testing Sample N: 8,110 8,245
Implementation

Email Received via CWIS

Establish ID of clients in KIDS and some relationships

Call comes in via County hotline

Generate

Risk Score

Button shown on KIDS
Service Elements...

What should happen?

How should the score be seen by the frontline social worker?
What should and **should not** happen when a call comes in?

*Depends on accuracy of the model vis-a-vis 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