Predictive Risk Modelling Adverse Outcomes for Children: Case Studies from New Zealand

Rhema Vaithianathan
Auckland University of Technology
Singapore Management University
Economists:
Rhema Vaithianathan, PhD
Professor
Department of Economics
AUT

Tim Maloney, PhD
Professor and Head Department
of Economics
AUT

Nan Jiang, PhD
Lecturer
Department of Economics
AUT

Child Protection:
Irene De Haan, PhD
Lecturer
Department of Social Work
University of Auckland

Claire Dale, PhD
Research Fellow
Department of Economics
University of Auckland

Emily Putnam-Hornstein, PhD
Assistant Professor
School of Social Work
University of Southern California

Ethics:
Tim Dare, PhD
Associate Professor
Department of Philosophy
University of Auckland

Ministry of Social Development
Targeting Preventive Care

Home visitation to..
  prevent maltreatment

Chronic disease management to..
  prevent re-hospitalisation
How well do health and social care systems choose children for preventive services?
Three Approaches

Clinical Judgement

Threshold Models

_Predictive Risk Modelling_
Threshold Models

Recruit people on the basis of a

*list* of characteristics

observed by a frontline *clinician*
Admissions criteria for a NZ home visiting program called “Family Start”

<table>
<thead>
<tr>
<th>LIST A. Family Start criteria</th>
<th>DESCRIPTION</th>
</tr>
</thead>
</table>
| Mental health issues                   | Either parent/carer has a mental health problem, for example:                                                                                                     • Post-natal depression  
• Anxiety  
• Depression  
• Self-harm or suicidal tendencies  
• Other (specify)  
Is this current or historic?  
Are they receiving any assistance or professional services for this?                                                                                                                                  |
| Difficulties with drugs, alcohol       | Either parent/carer has a problem with one of the following which affects parenting ability:                                                                                                          • Alcohol use  
• Illicit drug use  
• Gambling  
Are they are receiving any assistance from other professional services for this?                                                                                                                    |
| or gambling                             |                                                                                                                                                                                                                                                                            |
| Childhood history of abuse            | Either parent/carer experienced abuse as a child or young person.                                                                                                                                                                                                           |
Do they work?

• “FRET (an actuarial tool for ) is fairly useless as it always gives high or very high – they do not automatically open very high cases as they do not have the staff to do so”
  
  •  P Gillingham and Humphreys (2010)(interview subjects quoted on p 9).
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
What’s wrong with threshold models?

• Not developed or validated on the population
  – ODARA is used by the New Zealand police for family violence callouts but not validated on New Zealand population
  – LACE index is used in Ontario for readmission risk but found not to be valid for the NHS*

Automatic risk scoring tool which generates a risk score for an adverse event based on large administrative dataset.
Use of PRM in Health and Social Care

• Most advanced in the prediction of re-hospitalisation risk*

• Is known as “predictive analytics” – and routinely used in private firms

Stratified Approach to Prevention

Risk stratify

Access to services

P
R
M
High
Medium
Low
Example: Auckland hospitalisation PRM model (365 days readmission risk)

- When patients are discharged from hospital they are risk scored
- The score indicates the risk of re-hospitalisation within 365 days
- Risk score emailed to the Family Physician
- Case review high risk patients
- Currently being evaluated

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
Predicting Child Maltreatment in New Zealand
Can PRM be used to identify children (and pre-births) at high risk of a finding of maltreatment?

What are the practical issues to doing this?
Two Prototypes Developed

*Prototype I:* Using the Public Benefit system to identify children at risk* (~33% of a birth cohort)

*Prototype II:* “Using birth registry or public benefit system to identify children at risk** (~94% of a birth cohort by age 3 months, ~98% by age 6 months)

---


Prototype I: How it would work

A family registers on to the public benefit system

→ there is a <2-year-old child in the family
That night the system harvests Public benefit data set..
✓ Demographics of Primary Caregiver
✓ Demographics of Partner
✓ Previous benefits

Child Protection Data Sets...
Was Caregiver themselves abuse victim or known to CYF care and protection or youth justice services?
✓ Previous children with CYF contact?
✓ Previous children taken into care?
That **night** the system harvests

“SWIFTT” Data..

- Time since last prison spell of caregiver
- Caregiver is a refugee
- Foreign Born
Data

- Linkages between New Zealand’s public benefit and child protection systems
- Linked by the New Zealand Ministry of Social Development for the purposes of research
- Matched using a probabilistic algorithm that consisted of names, birth dates, and other personal identifiers
Cohort Used to Develop Algorithm

- Born between January 1, 2003, and June 1, 2006 and observed in the public benefit system between the start of the mother’s pregnancy and age 2.
- 103,397 public benefit spells, reflecting 57,986 unique children
- Used pseudonymised data – destroyed after use
Predictor Variables

The algorithm is a set of predictor variables and weights attached to each variable

We have a set of 224 predictor variables

We use stepwise probit to choose 132 variables

Develop model using 70% sample

Test on 30% sample
Using this model we can generate risk scores from 1 to 10 for each child ...
... predicts actual maltreatment rates well
Half of children in decile 10 will have maltreatment finding by age 5.
... 2% of children in decile 1
... never seen on a benefit have a rate of 1.4%
what proportion of finding will be “captured”?

(Based on risk score of the first spell)
Targeting the first spells in the top 2 deciles capture about

40 - 50 % of all findings that occur to children on the benefit
How long do we have

after they get the service offering

before there is a substantiated finding?
Being scored as in the top 20% and days to subsequent maltreatment finding

Percentage of children in top 20% of risk

Days to Substantiated Maltreatment Finding After Crossing top 20% risk Threshold

43%
Being scored as in the top 20% and days to subsequent maltreatment finding

Majority of maltreatment is more than 2 years after hitting the top 20% threshold
Prototype II

• Data from public benefit, care and protection, Corrections, birth and death registration and Ministry of Health administrative systems

• linkages formed by matching on names and dates of birth

• Apply to all children with a birth registration or included in a main welfare benefit within three months of birth (~94% of all new-born children)
Results

Suppose we target the top 5% of children in the 2007 birth cohort

... by age 5 $\Rightarrow$ 31% have maltreatment finding and 57% had 1 notification

... by age 10 (projected) $\Rightarrow$ 40% have a finding and 70% have a notification
### Prototype II -- Variables (base model)

<table>
<thead>
<tr>
<th><strong>Child</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Gender of child (male / female)</td>
</tr>
<tr>
<td>• Low birth weight or pre-term (yes / no or unknown)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Other children of the parents and/or caregivers</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Parenting demands (high demands - multiple birth child, other children under 2, or 3 plus other children / no other children / other children but not high demands)</td>
</tr>
<tr>
<td>• Other children with care and protection history in the last 5 years (yes / no)</td>
</tr>
</tbody>
</table>
# Prototype II -- Variables (base model)

## Parents and/or caregivers

- **Single** parent (yes / yes and father not listed on birth registration / no or unknown)
- At least one benefit caregiver is **not a birth registration parent** (yes / no / no birth registration by 3 months)
- **Age of mother** or primary benefit recipient when child was born (<20 / 20-24 / 25-29 / 30-34 / 35-39 /40+)
- Parents or caregivers with a **care and protection history** as a child (yes / no)
- Parents or caregivers with **findings of behavioural** or relationship difficulties as a child (yes / no)
- Mother or primary **benefit recipient's** time on benefit in the last 5 years (no time / 0-20% / 20-80% /80-100%)
- Parents’ or caregivers’ **mental health** in the last 5 years (no known substance abuse or other mental health disorder / substance abuse disorder / 3+ years in last 5 with substance abuse disorder / mental health disorder other than substance abuse / 3+ years in last 5 with mental health disorder other than substance abuse)
- Parents’ or caregivers’ **number of benefit address changes** in the last year (no address changes / 1-2 address changes / 3 plus address changes/missing - no benefit in last year)
- Parents’ or caregivers with a **Corrections history** in the last 5 years (no history / non-custodial sentence only / custodial sentence for non-violent crimes / custodial sentence for violent crimes)

## Community / office

- **CYF service centre** (43 categories)
Of the highest 5% of risk scored children in the 2007 sample...

<table>
<thead>
<tr>
<th>Category</th>
<th>% of all children</th>
<th>% of children with findings by age 2</th>
<th>% of the 3,000 children with the highest PRM scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male child</td>
<td>51.5</td>
<td>53.0</td>
<td>52.5</td>
</tr>
<tr>
<td>Low birth weight child or pre-term birth</td>
<td>8.2</td>
<td>10.7</td>
<td>8.2</td>
</tr>
<tr>
<td>Multiple birth child, other children under 2 or 3+ children</td>
<td>20.1</td>
<td>32.9</td>
<td>38.1</td>
</tr>
<tr>
<td>No other children (estimated)</td>
<td>59.4</td>
<td>51.1</td>
<td>45.7</td>
</tr>
<tr>
<td>Other children with a care and protection history in the last 5 years</td>
<td>4.8</td>
<td>34.9</td>
<td>59.9</td>
</tr>
<tr>
<td>Other children with a Police family violence notification or contact record in the last year</td>
<td>0.8</td>
<td>8.6</td>
<td>12.4</td>
</tr>
<tr>
<td>Single parent</td>
<td>24.6</td>
<td>74.3</td>
<td>87.7</td>
</tr>
<tr>
<td>No birth registration at 3 months of age</td>
<td>6.3</td>
<td>26.1</td>
<td>35.5</td>
</tr>
<tr>
<td>Mother or caregiver aged under 25</td>
<td>25.3</td>
<td>53.5</td>
<td>54.6</td>
</tr>
<tr>
<td>Parents or caregivers with a care and protection history as a child</td>
<td>9.8</td>
<td>43.2</td>
<td>57.2</td>
</tr>
<tr>
<td>Parents or caregivers with findings of behavioural or relationship difficulties as a child</td>
<td>3.5</td>
<td>19.1</td>
<td>29.1</td>
</tr>
<tr>
<td>Child seen on benefit by 3 months of age</td>
<td>23.2</td>
<td>79.0</td>
<td>92.3</td>
</tr>
<tr>
<td>Mother or caregiver's spent 80-100% of time on benefit in the last 5 years</td>
<td>11.7</td>
<td>50.7</td>
<td>65.5</td>
</tr>
<tr>
<td>Parents or caregivers received benefit for a substance abuse disorder in the last 5 years</td>
<td>1.5</td>
<td>8.2</td>
<td>14.1</td>
</tr>
<tr>
<td>Parents or caregivers received benefit for other mental health disorder in the last 5 years</td>
<td>4.9</td>
<td>17.4</td>
<td>19.0</td>
</tr>
<tr>
<td>Primary caregiver's with 1+ address changes recorded in benefit data in the last year</td>
<td>11.6</td>
<td>26.1</td>
<td>30.4</td>
</tr>
<tr>
<td>Parents or caregivers with any sentence in the last 5 years</td>
<td>7.2</td>
<td>25.5</td>
<td>33.6</td>
</tr>
<tr>
<td>Parents or caregivers with a custodial sentence in the last 5 years</td>
<td>3.1</td>
<td>11.1</td>
<td>16.6</td>
</tr>
<tr>
<td>High deprivation neighborhood (deciles 8-10, of non-missing)</td>
<td>36.9</td>
<td>69.0</td>
<td>70.7</td>
</tr>
</tbody>
</table>
Coincidence of Family Start Threshold and PRM

What percentage of children identified as high risk by PRM were identified by existing threshold models?

- 7% of all births were included in Family Start
- But... only 56% of them were in the highest risk score decile (according to the PRM)
- ( Might be due to refusal to participate)
What do you need to undertake a PRM?

- Rich data – does not need to be matched in real time
- Point at which clients come into contact with this data system is “early” in the trajectory
- Priority lists can be generated from this data
- Non-stigmatising interventions are available
- Collaboration between public agencies, multi-disciplinary researchers and legislators
Further reading


