Using Machine Learning to Improve Child Welfare Decision Making in Allegheny County, PA

Dennis P. Culhane

Source: Erin Dalton, Allegheny County, Department of Human Services
What is an IDS?
Child Welfare Case Study: Allegheny County, PA
The Problem: Ineffective Screening

- 4 in 5 children in Allegheny County who died (or nearly died) as a result of abuse were reported but not screened in for investigation.
The Proposed Solution: Risk Scoring

• Use predictive modeling to generate risk scores
• Call center would use risk score to inform investigation decision
Using Predictive Analytics to Prevent Child Abuse

- Issued an RFP
- 16 organizations responded
- Many multidisciplinary teams
- Awarded—Auckland University
  - Experience
  - Concern with ethical issues
  - Implementation support
What data go into the model?

- Mental Health
- County Prison
- Drug & Alcohol
- Juvenile Justice
- Probation
- Public Welfare
- Parent History
- Community Indicators
- Child & Family History
How the Score is Generated

KIDS Application

Call Screening Module:
- Call Info (Caller info, client Info)
- Client MCI Clearance/Creation
- Allegations
- Relationships (not mandatory, but last process requires it)
- Collatorals, other non-mandatory screens
- Screening Outcome

Intake → Etc...

Risk score saved, per run on MCI ID, Run ID, with Algorithm version #

Algorithm

Includes logic to: pull from DW and KIDS staging data + rules to calculate 289 base algorithm variables + user-defined variables

KIDS DB

Automatically run Algorithm, per MCI ID, after Clearance and after Relationships

Return Risk score per MCI, Run ID, Run date/time, Algorithm version # to DB only

Algorithm Configuration Application

Algorithm Configuration Settings
Version: [34]
Eff Begin Dte: 8/1/15
Eff End Dte:

Var 1 | Active? | Weight
Var 2 | Active? | Weight
Var 3 | Active? | Weight
Var N | Active? | Weight

Search Algorithm Configuration
Search By Effective Date or Version # to view previous algorithm configurations

User-Defined Variables

Per MCI ID, pull staged data from KIDS + DW, and pull real-time data from KIDS DB.

200 calculated variables archived

For every MCI ID Risk Score run, archive all of the input data feed into the algorithm

Image Source: Allegheny County DHS
The screening score is from 1 to 20

The higher the score, the higher the chance of the future event (e.g., abuse, placement, re-referral) according to the data
Testing the Screening Model

• Tested model’s accuracy using thousands of historical maltreatment calls
• Followed the children in subsequent referrals to see how often the model was correct...
Results: Predicting Re-referrals

- **Risk Score of 1**: 1 in 10 re-referred within 2 years of a call
- **Risk Score of 20**: 9 in 10 re-referred within 2 years of a call
Results: Predicting Out-of-Home Placement

• 1 in 100 children who received a score of 1 were placed in out-of-home care within 2 years of the call

• 1 in 2 children who received a risk score of 20 were placed in out-of-home care within 2 years of the call
Without the predictive model...

• 27% of the highest risk cases were screened out
• 48% of lowest risk cases were screened in
Model Validation — Preliminary Results

- Used Children’s Hospital injury data
- Of all children referred to child welfare between 2010-2016, 21% had a hospital admission between 2002-2015
  - 5% of those kids had a Children’s Hospital code for external injury
- Risk score of 20: 45% of those kids had a hospitalization & 16% had an external injury code
- Risk score of 1: 9% are hospitalized, 2% had an external injury code
Addressing Ethical Issues

• Multiple community meetings
• Independent ethical review
• Scores *never* generated based only on demographic information. Child and parent must have positive MCI ID.
There is a real worry that *not* using this technology is unethical...It’s giving us information to make a difference in kids’ lives.

Tim Dare, Professor of Philosophy, University of Auckland
Operationalizing Results

• Calls are now assessed using the risk score and information gathered from the caller.

• All cases that score a **16 or higher** are **automatically referred** to investigation. Otherwise, the score is supplemental information.
Final Thoughts & Next Steps

• Existence of the IDS and 20+ years of successful IDS use made this an easier implementation in Allegheny County.
• Multiple jurisdictions are planning to replicate.
• Allegheny County now working on using predictive modeling to address other child welfare issues.
• Will soon release methodology paper & ethical analysis results.