

Date: 24 September 2013

Subject: Final External Peer Review for the New Zealand Ministry of Social Development's PRM Development Project

Submitted by: David R. Schwartz

Title: Director, Research and Evaluation at the William Penn Foundation

After reviewing two iterations of predictive risk model (PRM) development and testing, it is clear that the Ministry has leveraged exceptional in-house statistical modeling capacity to develop a cluster of relatively accurate models that could potentially improve prevention decisions made on behalf of vulnerable children and families. While the trained PRM algorithms ultimately lacked the accuracy needed to initiate more coercive child protective services, the algorithms were able to accurately identify a cohort of children and their families that could potentially benefit from an array of supportive social and health services. The Ministry has wisely recommended that, in subsequent experiments, PRM algorithms should strictly be used as decision support for professional social workers in the field. The prevention/early warning system focus of the PRM (reading from birth records prior to child protective services involvement) could eventually provide social workers and other professional colleagues with a longer runway for helping children and families before it is too late. It is encouraging to see that the New Zealand Ministry of Social Development has taken such a robust look at their data. This exciting work has the potential to be a bellwether approach for practice and assessment in child welfare.

The Ministry is aware of the fact that a lot of hard work still needs to be done to ensure that PRM systems are implemented carefully, for example:

- Model Optimization and Testing
 - Is one model ideal, or is an ensemble of models more accurate given different subpopulations?
 - When different models are included for training and testing, does the Ministry have the capacity to optimize each of the models to ensure that there is a fair comparison?
 - Should the Ministry look to PRMs that are easier to interpret (e.g., PRMs that grow easy to understand trees)?
 - How often should the models be refreshed and trained/tested with new data?
- Practitioner Involvement
 - Should practitioners be exposed to the rule sets under the hood of the PRM?
 - Should practitioners be involved in the ongoing development of PRMs?
 - Should practitioners be engaged in the designing of practitioner-facing interfaces to PRM?
 - If a model is slightly less accurate than the most accurate model, yet provides practitioners with understandable rule sets/decision trees, which PRM should be selected?

- What is the most efficient and effective method for training workers to use PRM tools to inform their practice decisions?
- Scope of PRM
 - Should the PRM be trained to intelligently match children and their families with the services that give them the best chance at successful outcomes?
 - Should PRMs be developed to help supervisors identify patterns in child welfare worker performance data?

Given all of the effort the Ministry has put into developing this promising work, the next logical step would seem to be carefully experimenting with PRMs as a decision support tool. If the decision is made to utilize a randomized controlled trial design to further test the PRM and its implications for practice, it is recommended that the Ministry complement the RCT with rigorous qualitative research and clinical data mining (e.g., are there typologies of workers that do not respond well to PRM?).

It is important to mention that the Ministry has exhibited extraordinary capacity in the following areas:

- Advanced research project management
- Data integration
- Data preparation
- Predictive analytic modeling
- Model testing
- Research synthesis
- Translating research into practice

It is unclear if other child welfare agencies internationally have the combination of personnel, technology, and data it takes to develop a carefully designed and tested PRM project in-house. The Ministry's cutting edge work in this area could significantly influence child welfare system staffing, data integration, and assessment tool decisions for many years to come.