Predictive Modeling

Potential Application in Prevention Services

Moira L. Wilson, BA (Hons), Sarah Tumen, PhD, Rissa Ota, PhD, Anthony G. Simmers, BSc (Hons)

Introduction: In 2012, the New Zealand Government announced a proposal to introduce predictive risk models (PRMs) to help professionals identify and assess children at risk of abuse or neglect as part of a preventive early intervention strategy, subject to further feasibility study and trialing. The purpose of this study is to examine technical feasibility and predictive validity of the proposal, focusing on a PRM that would draw on population-wide linked administrative data to identify newborn children who are at high priority for intensive preventive services.

Methods: Data analysis was conducted in 2013 based on data collected in 2000–2012. A PRM was developed using data for children born in 2010 and externally validated for children born in 2007, examining outcomes to age 5 years.

Results: Performance of the PRM in predicting administratively recorded substantiations of maltreatment was good compared to the performance of other tools reviewed in the literature, both overall, and for indigenous Māori children.

Conclusions: Some, but not all, of the children who go on to have recorded substantiations of maltreatment could be identified early using PRMs. PRMs should be considered as a potential complement to, rather than a replacement for, professional judgment. Trials are needed to establish whether risks can be mitigated and PRMs can make a positive contribution to frontline practice, engagement in preventive services, and outcomes for children. Deciding whether to proceed to trial requires balancing a range of considerations, including ethical and privacy risks and the risk of compounding surveillance bias.

(Am J Prev Med 2015;48(5):509–519) Crown Copyright © 2015 Published by American Journal of Preventive Medicine All rights reserved.

Introduction

The New Zealand Government's 2012 *White Paper for Vulnerable Children*[1] includes a proposal to use a predictive risk model (PRM) to help professionals identify and assess children at risk of abuse or neglect as part of a preventive early intervention strategy. The white paper acknowledged that PRMs appear promising based on prototype research,[2] but carry ethical risks and warrant careful feasibility study and trialing. Decisions are yet to be made on whether to proceed to trial, and if so the application of PRM that would be trialed (e.g., use in identifying children for referral to voluntary services as outlined in the white paper, or use to support service prioritization for children already referred to services). Within social services, use of PRMs based on administrative data is most advanced in health care.[3] Although greater utilization of existing administrative data and application of advanced computational techniques in child welfare have been explored and advocated for,[4–7] use of PRMs to support child maltreatment prevention efforts has received less attention.[2,7]

This paper presents findings from a feasibility study building on prototype research. The purpose of the study was to examine technical feasibility and predictive validity of the white paper proposal, focusing on a PRM that would identify newborn children who are at high priority for intensive preventive services. Where prototype research considered risk among young children in the public benefit system,[2] the present study examined feasibility of a population-wide PRM. Parallel studies considered the ethics of the white paper proposal

---

From the Insights MSD (Wilson, Ota, Simmers), Ministry of Social Development; and Analytics and Insights, Performance Hub (Tumen), New Zealand Treasury, Wellington, New Zealand

Address correspondence to: Moira L. Wilson, BA (Hons), Principal Analyst, Insights MSD, Ministry of Social Development, P.O. Box 1556, Wellington 6140, New Zealand. E-mail: moira.wilson001@msd.govt.nz.

0749-3797/$36.00
http://dx.doi.org/10.1016/j.amepre.2014.12.003

Crown Copyright © 2015 Published by American Journal of Preventive Medicine All rights reserved.

(T. Dare, University of Auckland, unpublished observations, 2013), and implications for indigenous Māori (A. Blank, unpublished observations, 2013).

Methods

Study Sample

Data analysis was conducted in 2013. The study linked data collected in 2000–2012, including (1) birth information held by the Registrar-General of Births, Deaths and Marriages; (2) benefit data for the child, other children in the family, and their parents or caregivers held by the Ministry of Social Development (MSD); (3) MSD Child Youth and Family (CYF) child welfare data for the child, other children in the family, and their parents or caregivers (relating to their own childhood); (4) data on sentences administered by the Department of Corrections served by the parents or caregivers; and (5) Ministry of Health data for the mother, child, and recently born siblings, including administrative markers of maternal mental health and transience, and sibling histories of intentional injury hospitalizations. Reported findings are from models that excluded Ministry of Health data, as these were found to not improve predictive accuracy (MSD, unpublished observations, 2014). Ethics approval was granted by the Central Region Health and Disability Ethics Committee, and affirmed by the National Ethics Advisory Committee.[8]

It was necessary to link data using names and dates of birth, given that there was no unique identifier across different administrative systems. Linkage of data was conservative in that data were only linked if there was a very high level of agreement among identity details. The aim was to approximate an administrative approach with very low tolerance for false-positive data linkages. Prototype research was based on a less conservative linkage.[2] Clerical checking undertaken for this study showed that compared to the conservative approach to linkage, the less conservative approach resulted in fewer missed matches but generated more false-positive data linkages.[9]

The study population included all live-born children who were identified within 91 days of birth as a result of either (1) a birth registration being filed or (2) inclusion in a main public benefit (this covered an estimated 94% of all New Zealand live-born children and yielded records for approximately 60,000 children in each annual birth cohort).

Two data sets were created: (1) a sample data set for study population children born in 2010 was created to develop the model and for internal validation, and (2) a data set that contained data for all study population children born in 2007 was created for external model validation. The predictive accuracy and model performance statistics reported here are based on external validation tests only. Results of internal validation were similar to external validation tests. The 2010 sample data set included all children born in that year who had findings of maltreatment by age 2 years, and the ratio between event and no event groups in the sample was kept at 1/5. The sample data set was partitioned using the 70/30 rule, where a training data set contained 70% of records that were used for model estimation and the model was internally validated on the remaining 30% of records. A stratified partitioning method was used to ensure both the training and validation data sets had the same proportion of events. Tests suggested very little sensitivity to sampling method, data partitioning rules, or birth cohort used for model estimation.

Measures

The targeted outcome variable was whether the child had by age 2 years at least one substantiated finding of maltreatment (emotional, physical, or sexual abuse or neglect—acknowledging that this likely understates the proportion of the cohort that actually experienced maltreatment by this age).[10] A range of other child welfare outcomes were considered and tested as target variables. Although models for most performed well, models predicting substantiated physical abuse did not.

Predictor variables were indicators of risk and protection derived from administrative data (Table 1). Variable selection was informed by the literature on risk and protective factors for abuse and neglect,[10–12] and by previous studies examining associations between administrative data available at birth and subsequent reporting or substantiation of maltreatment.[4,6] This approach to predictor variable selection was chosen for transparency and ease of translation. The time the caregiver had spent supported by welfare benefits was included as a proxy for poverty and its persistence. In New Zealand, benefit receipt has strong associations with child poverty and measured material deprivation.[13,14]

Statistical Analysis

Twelve modeling algorithms were tested and compared, including multilevel modeling to account for the influence of local context.[15,16] Most had similar predictive performance (Table 2). Stepwise logistic regression based on the maximum likelihood estimator method was selected as the preferred modeling strategy because it allowed researchers to gain knowledge about the most useful predictors. The significance entry level for predictor variables was set to $p < 1.00$, allowing all variables to enter the model. The significance stay level was set to $p < 0.02$, allowing variables to remain in the model only if their significance was less than $p < 0.02$ when the effect of other variables was controlled.

Results

Thirteen of 15 variables considered satisfied the stepwise selection criteria and remained in the model. The top three predictors selected were as follows: (1) the presence of other children with contact with care and protection services in the previous 5 years; (2) the length of time the mother or primary benefit recipient with care of the subject child was supported by benefit in the previous 5 years; and (3) having a parent or caregiver who was known to have had contact with care and protection services in their own childhood. Parameter estimates and ORs are provided in Table 3. Sensitivity tests showed considerable stability in the top predictors.

A perfect prediction tool would correctly identify every child who would subsequently go on to experience the outcome of concern and not incorrectly identify any child

Download English Version:

https://daneshyari.com/en/article/6237781

Download Persian Version:

https://daneshyari.com/article/6237781

Daneshyari.com