MINISTRY OF
SOCIAL DEVELOPMENT
Te Manatū Whakahiato Ora

Date: 2 February 2014

The feasibility of using predictive risk modelling to identify new-born children who are high priority for preventive services[1]

In response to a proposal in the White Paper for Vulnerable Children, this study assesses whether predictive risk modelling (PRM) tools that draw on linked administrative data can be used in practice to identify early new-born children at high risk of maltreatment as part of a strategy to prevent maltreatment from occurring. The feasibility of linking administrative data to support PRM is assessed, and the predictive accuracy of PRM models drawing on different agency data and targeting different measures is compared.

We find that linkage of administrative data to support PRM is feasible, but the linkage is subject to error and a system for review would be needed in any implementation. Assessment of the true predictive accuracy of the PRM models developed is not possible because much of the abuse and neglect that occurs goes undetected. However, performance of the models in predicting administratively recorded substantiations of maltreatment is good compared to other tools reviewed in the international literature, both overall, and for Māori children. Some but not all children who go on to have contact with care and protection services or to experience substantiated maltreatment can be identified using this approach. We recommend that, if taken to trial, PRM tools should not be the sole mechanism for identification and referral of children at high risk, and should be used in combination with professional judgement.

Findings from this feasibility study, in combination with those from an ethical review, will inform decisions around next steps in relation to the possible trialling of PRM.

TABLE OF CONTENTS

Executive summary

The New Zealand Government's 2012 White Paper for Vulnerable Children includes a proposal to use predictive risk modelling (PRM) tools to assist professionals in identifying which children are at risk of abuse or neglect to support a preventive early intervention strategy, subject to the outcomes of feasibility study and trialling.

Under the proposal, PRM tools would draw on existing administrative data to provide an estimate of the risk of future maltreatment. Where the risk score was above an agreed threshold, the child's details would be passed to a local Children's Team. Depending on the assessment of the Children's Team, this might lead to out-reach to the families of the children identified, and the offer of support and services.

The White Paper proposed that Children's Teams would target children whose level of risk is just below that which would require a statutory care and protection response. The aim would be to prevent these children from requiring statutory services later on. It was proposed that PRM tools would not be the only source of referral to Children's Teams. Referrals would also be made by front-line professionals (New Zealand Government, 2012).

Decision making to date in the New Zealand care and protection system and social services has generally relied on consensus-based risk screening models and clinical judgement, with broad referral guides in use in some areas, and limited use of standardised "actuarial" tools that obtain a risk score from a checklist or questionnaire. This would be the first application of PRM tools.

Internationally, there has been very limited application of advanced computational tools to support decision making in child protection or early intervention, although actuarial tools are in use in many jurisdictions. Use of PRM for early identification of risk of child maltreatment is untried.

The White Paper acknowledged that PRM appears promising based on preliminary research undertaken by the University of Auckland (Vaithianathan et al., 2012), but carries ethical risks, and warrants careful, staged, feasibility study and trialling.

This report presents findings from a study that builds on the preliminary research. It examines the feasibility of a population-wide PRM tool for identifying new-born children who are at high risk of future maltreatment for referral to Children's Teams, and considers:

- technical feasibility (including the feasibility of data linkage, the predictive accuracy that could be achieved and the modelling strategy, agency data and predictor variables that would support the best predictions of risk)
- predictive accuracy specifically for Māori children who are over-represented among those for whom there are substantiated findings of maltreatment, and the proportionality of predictions for Māori children relative to their share of known maltreatment
- proportionality of predictions relative to known maltreatment for other social groups
- legal and operational feasibility and the lags associated with accessing selected agency data that would impact on the timeliness of a PRM tool.

Ethics approval for the study was granted by the Central Region Health and Disability Ethics Committee, a decision affirmed by the National Ethics Advisory Committee.

The feasibility study has proceeded in parallel with an ethical review of use of PRM in this context to identify ethical concerns and propose resolutions.

The ethical review traverses a range of issues, including trade-offs between the rights of children to protection from harm and the right to privacy, the need to avoid stigma, the use of data without consent, concerns related to resource allocation, ethical concerns from a Māori perspective, and the obligation of agencies to act in the face of elevated risk scores.

Feasibility of data linkage

Administrative data from benefit, care and protection, Corrections, birth and death registration and Ministry of Health administrative systems were linked for the purposes of the feasibility study.

With no unique identifier against which all of an individual's contact with different government agencies is recorded, linkages must be formed by matching on details such as names and dates of birth. This linkage of data to support PRM is feasible, but is subject to error.

The linkage of data used in this study is "conservative" in that data from two agencies are only linked if there is a very high level of agreement between identity details in the two systems. Linkage of data for a child occurs only if the given name of one of the parents or caregivers is the same, in order to confirm the match.

The aim is to approximate an administrative approach which would seek to minimise the likelihood of erroneously linking one individual's information to another individual. A check of a random sample of linkages shows that this conservative approach does lead to some instances in which information that should inform the assessment of risk for a child is missed, but found no cases of information for different people being linked.

Care and protection data present particular challenges for data linkage as care and protection records for a child are often established without identity verification.

Applying a less conservative approach to data linkages reduces the rate at which linkages are missed and improves measures of predictive performance, but also introduces some false-positive linkages where one individual's information is erroneously linked to another individual.

Given the potential for error in data linkage, it is recommended that there be a system for review of the linkages that inform the risk scoring before details of a high risk child are used as part of any working PRM. Having an effective system of review in place may allow a less conservative approach to data linkage, and improvements in predictive accuracy to be realised.

Feasibility of a population-wide PRM tool that would apply around the time of birth

PRM tools use relationships in existing administrative data to estimate the likelihood of a future event.

To assess the feasibility and predictive performance of a population-wide PRM tool for identifying new-born children who are high priority for referral to a Children's Team, models were developed using conservatively linked research data for recent birth cohorts.

These models would apply to all children with a birth registration or included in a main welfare benefit within three months of birth (covering an estimated 94 percent of all new-born children) and would predict the risk of findings of substantiated maltreatment of any type (emotional, physical or sexual abuse or neglect) by age two.

The base models draw on predictor variables or risk markers drawn from birth, benefit, care and protection and sentencing data.

Preliminary research established that ethnic group does not need to be included as a predictor in order to support PRM, adding only marginally to predictive performance.

Assessment of the true predictive accuracy of the models developed is not possible because much of the abuse and neglect that occurs goes undetected.

The models developed perform well in predicting risk of agency contact. To illustrate predictive accuracy, if the five percent of children in the 2007 birth cohort with the highest predicted risk scores based on a PRM tool had been considered high priority for referral to Children's Teams:

- by age five, 31 percent of the children referred would, in the absence of any additional intervention, have had a substantiated finding of maltreatment and 57 percent would have had at least one notification

- projecting forwards to age ten, the proportion with a substantiated finding of maltreatment is expected to increase to four in ten and the proportion who are notified is expected to rise to seven in ten

- the referred group would have accounted for 32 percent of *all the children* who had substantiated findings of maltreatment by age five and 36 percent of *all maltreatment findings* (counting events rather than children)

- the referred group would have accounted for 22 percent of all the children who had a notification by age five and 31 percent of *all notifications* to that age

- in 96 percent of cases where a child had no findings of maltreatment by age 5, they would not have been referred

- 2,163 of the 3,114 children referred (69 percent) would have no substantiated finding of maltreatment by age five (false positives).

Based on the ethnicities of the child recorded in birth registration data, positive predictive value and sensitivity for Māori children is better than that found overall.

Implementation of the PRM proposal outlined in the White Paper would require a decision to be made about the threshold for referral. This decision would need to be informed by a range of factors, including the capacity of the Children's Teams and the availability of services, as well as predictive accuracy.

Choice of threshold involves a trade-off. Tightly targeting to those at the highest risk means including only a small proportion of the children who would go on to be maltreated in the referred group. Including a larger proportion of the children who would go on to be maltreated in the referred group would mean targeting a group at lower risk, on average.

Regardless of the threshold chosen the output from a PRM is a probability of a future event. By its very nature, a high risk score may suggest that the event is likely, but it does not suggest it is sure to happen. Models will identify as low priority some children who are subsequently the victims of high profile abuse or neglect, and identify as high priority children who never experience maltreatment. This needs to be acknowledged as an inevitable limitation of any early identification approach, including both PRM and human decision making.

Comparison with other tools for early identification in reviews of the international literature

The available evidence is that actuarial assessment tools based on empirical evidence are more accurate than human decision making (consensus-based models or professional judgment) in the assessment of risk of future harm to children.

At the same time, evidence reviews consistently find limited accuracy in predicting child maltreatment from the small number of tools that have been developed with the aim of predicting child maltreatment or ensuring that preventive early intervention services are targeted to families most in need.

At selected thresholds, the PRM tools examined would meet the criteria of a suggested guide to what should be able to be achieved with a "good" predictor for child maltreatment. In a review of the international literature, only one of the previously developed tools reviewed could meet these criteria.

Proportionality of predictions relative to known maltreatment

The models developed would refer children in numbers that do not always reflect their share of known maltreatment.

Māori children would be slightly over-represented among the children referred, relative to their share of known maltreatment (comprising 69 percent of the 3,000 children with the highest risk scores compared with 61 percent of children with findings of maltreatment by age two, where ethnicity is known).

In addition, relative to their share of known maltreatment, the base models would refer too many children:

- with parents or caregivers who receive benefit (comprising 92 percent of the 3,000 children with the highest risk scores compared with 79 percent of children known to be maltreated by age two)

- in a family where other children have had contact with care and protection services (60 percent compared with 35 percent)

- with parents or caregivers with known childhood contact with care and protection services (57 percent compared with 43 percent).

Over-representation of Māori children relative to their share of known maltreatment can be addressed by deploying separate models for Māori children and children who are not known to be Māori and selecting the highest risk scored children for referral from each model in proportion to shares of known maltreatment. Test models show that this can be done without any notable loss of predictive accuracy.

The same approach can be used to address disproportionality on other dimensions. Tests combining separate models by benefit status and prior care and protection history show little or no loss of predictive accuracy. However, a solution that resolves disproportionality on all dimensions of concern is unlikely.

Alternative target variables

Models targeting a range of different care and protection outcomes perform well and could be explored further in any operationalisation.

However, due to poorer model performance and the rarity of the events targeted, development of a model for predicting which children will have substantiated findings of physical abuse or will be hospitalised for maltreatment related injuries or marker injuries is not feasible.

Agency data required

Compared to a PRM tool that draws on benefit and care and protection data only, and screens only the population of children supported by main benefits:

- adding Corrections sentencing data provides no substantive gain over and above information on imprisonment that can be obtained from benefit data

- adding births data provides no substantive gain in predictive accuracy, but allows screening across the population, and goes some way to addressing the disproportionate representation of children supported by benefits among those identified as at high risk, and provides a basis for the deployment of separate models to address disproportionate representation of sub-groups including Māori children

- beyond the incorporation of births data, the further addition of Ministry of Health data provides very little gain in predictive accuracy, and only modest benefit in reducing disproportionality on selected dimensions.

Main predictors

The main predictors are relatively stable across the models developed. They include:

- the presence of previous children with contact with care and protection services in the last five years

- the length of time the parent or caregiver was supported by main benefits in the last five years

- having a parent or caregiver who had contact with care and protection services in their own childhood.

Other variables with high predictive utility include indicators related to mental health, location, sentencing history, family violence, single parent status and caregiver age.

These cannot be interpreted as factors that cause maltreatment to occur, only as useful predictors or risk markers for identifying some of the children at high risk.

Legal and operational feasibility

Information sharing and data linkage for a fully implemented PRM, where it involves linking data held by different Government agencies, may require new authorisation.

Operationalising a PRM would require systems that ensure the secure handling of the data drawn on and generated, and for ensuring protection from mis-use.

 A Privacy Impact Assessment would be a useful first step in any implementation. It would provide a framework for privacy protection in systems that feed in data from contributing agencies and databases, that link data, that apply and regularly review risk scoring algorithms and that generate and disseminate PRM information.

Conclusions

PRM tools based on linked administrative data can be used to identify early some of the new-born children at high risk of maltreatment. Compared to other tools developed to predict the risk of future maltreatment for new-born children and reviewed in the international literature, these models have good predictive accuracy, both overall, and for Māori children.

Not all children who go on to experience maltreatment are able to be identified early using this approach. We recommend that, if taken to trial, PRM tools should therefore not be the sole mechanism for identification and referral of children at high risk.

Careful thought would need to be given to other potential uses of predictive risk modelling information, and the role this information plays in risk assessment and decision making. The

administrative data can only provide partial and selective capture of risk information and are subject to error in linkage. A cautious approach should therefore be taken to making the scores, or the information that sits behind those scores, available to front-line professionals.

Next steps

Findings from this feasibility study, in combination with those from the ethical review, will inform decisions around next steps in relation to the possible trialling of PRM.

Beyond this study, the data assembled provide the basis for a wide range of new analysis that can aid the understanding of, and response to, child maltreatment.

Predictions from the PRM models developed in this study could be compared with those from models that quantify and score according to the accumulation of risk factors. New ways of describing the data (for example using cluster analysis techniques to profile sub-groups of children who are known to experience maltreatment or are identified as high risk), could also be explored.

Research is under way to better understand Māori children's disproportionate representation in maltreatment statistics, and to explore more generally whether bias exists in measured contact with care and protection services.

Finally, with further data linkage, the data offer new opportunities for analysis to build evidence on the impacts of existing services on outcomes for children.

1. Introduction

1   The purpose of this report is to present findings from a study that examines the feasibility of a predictive risk modelling (PRM) tool that would identify new-born children who are at high risk of future maltreatment based on existing administrative data.

2   The study considers:

- technical feasibility (including the feasibilty of data linkage, the predictive accuracy that could be achieved and the modelling strategies, agency data and predictor variables that would support the best predictions of risk)

- predictive accuracy specifically for Māori children who are over-represented among those for whom there are substantiated findings of maltreatment, and the proportionality of predictions for Māori children relative to their share of known maltreatment

- proportionality of predictions relative to known maltreatment for other social groups

- legal and operational feasibility of selected data feeds and the lags associated with each.

3   A technical companion report provides more details of variable definition and the modelling strategy and sensitivity testing.

4   A separate stream of work has reviewed the ethics of use of PRM tools in this context (Dare, 2013) and considered ethical concerns from a Māori perspective (Blank et al., forthcoming).

5   Together, the ethical review reports and the findings from this feasibility study will inform the development of advice on whether and how PRM might be able to be applied in ways that mitigate or minimise ethical risks, and whether and how to take these options to trial.

Outline

6   Section 2 of this report provides the background to the feasibility study and a brief review of existing international literature on tools for early identification of children at high risk of future maltreatment. The report then introduces a "base model" that demonstrates the PRM approach (section 3), and details its predictive performance overall and for Māori children (section 4). Section 5 considers the accuracy of the underlying data linkage and the sensitivity of the results to the data linkage approach taken.

7   The proportionality of the predictions relative to the share of known maltreatment that is accounted for by Māori children and other social groups is then considered (section 6). Sections that follow test sensitivity to the use of separate models to control disproportionality (section 7), the definition of the target variable - acknowledging that substantiated maltreatment may be an imperfect proxy for harm occurring (section 8), and the agency data that is drawn on, including consideration of whether inclusion of Ministry of Health data would improve predictions (section 9).

8   Subsequent sections consider the most useful predictors (section 10), time lags in the flows of selected agency administrative data (section 11), legal and technical feasibility (section 12) and the limitations of the study (section 13), before discussing the findings (section 14), presenting conclusions (section 15), and outlining next steps (section 16).

2. Background

9   The New Zealand Government's White Paper for Vulnerable Children includes a range of proposals aimed at preventing maltreatment. One of the proposals is to use PRM tools to assist professionals in identifying and assessing which children are at risk of abuse or neglect, subject to the outcomes of feasibility study and trialling (New Zealand Government, 2012, pp.79-81).

10   PRM tools would draw on existing social sector administrative data. Under the White Paper proposal, where the tools estimated the risk of future harm for a child was above an agreed threshold, the child's details would be passed to a local Children's Team comprised of front-line professionals working with families and children. This would sit alongside other pathways for referral to a Children's Team.

11   Children's Teams will work with children whose risk level is just below that which would require a statutory care and protection response. Those above the threshold will continue to be served by care and protection services. Depending on the assessment of the Children's Team, referral by a PRM might lead to out-reach to the families of the identified children by a known professional (such as a midwife or Well Child/Tamariki Ora provider), and the offer of support and services aimed at preventing maltreatment from occurring.

12   PRM would not be the only pathway for referral to a Children's Team. Children not identified by PRM would still be able to be identified as at risk and referred for assessment and services by frontline professionals, including care and protection social workers (New Zealand Government, 2012).

13   To date, risk assessment in the New Zealand care and protection system and social services has generally relied on consensus-based risk screening models and clinical judgement, with broad referral guides in use in some areas,[2] and limited use of standardised "actuarial" tools.[3] This would be the first application of PRM tools.

14   The White Paper PRM proposal was informed by preliminary research undertaken by the University of Auckland and completed in 2012 which found that PRM based on linked benefit and care and protection data held promise as a means of prospectively identifying children at high risk of going on to have a substantiated finding of maltreatment (Vaithianathan et al., 2012; Vaithianathan et al., 2013).

15   The University of Auckland researchers recommended a full ethical evaluation of PRM and the development of an ethical framework to guide agencies in their responses to the use of risk scores before any implementation. They highlighted ethical concerns about applying PRM only to children known to the benefit system and recommended broadening the research data in order to include all children in the population. They also recommended including health data and community-level characteristics as predictors (Vaithianathan et al., 2012).

16   The feasibility study reported on here responds to the recommendations made in that report. It investigates the feasibility, data requirements and predictive power of population-wide PRM tools, focussing on PRM that would apply to all New Zealand-born children at or around the time of birth.

---

[2] Family Start referral criteria for example http://www.familyservices.govt.nz/documents/working-with-us/programmes-services/early-intervention/new-family-start-march-2012/family-start-referral-guide.pdf

[3] The Canadian ODARA (Ontario Domestic Assault Risk Assessment) tool is used by Police in New Zealand to predict the likelihood of re-assault following family violence incidents (Curtis, 2012).

17 Ethics approval for the feasibility study, and in particular for the incorporation of health data, was granted by the Central Region Health and Disability Ethics Committee on 23 November 2012 (Ethics Ref 12/CEN/46), a decision affirmed by the National Ethics Advisory Committee in February 2013. A Privacy Impact Assessment for the research was discussed with the Office of the Privacy Commissioner on 29 November 2012.[4]

18 The White Paper acknowledged that PRM appears promising based on preliminary research, but is untried in the context of child maltreatment, carries ethical risks, and warrants careful, staged, development (New Zealand Government, 2012).

19 Other projects planned as part of this staged development include qualitative interviews with front-line professionals in Children's Team pilot sites to inform the design of options that could be considered for trialling, and a trial design for assessing PRM's efficacy and cost effectiveness in helping prevent maltreatment.

Existing literature on tools for early identification

20 Decision making in child protection is intrinsically difficult, uncertain, and subject to the practical limits to our ability to predict future outcomes (Munro, 2011; Mansell et al., 2011; Leventhal, 1988). Much of the abuse and neglect that occurs goes undetected (Munro, 2011; Gilbert et al., 2009; Runyan et al., 2002), and risks can be concealed. This presents a fundamental difficulty for risk prediction as it can only be informed by relationships between observed risk and protective factors and the subset of maltreatment that is known to occur.

21 The available evidence is that actuarial assessment tools based on empirical evidence are more accurate than human decision making (consensus-based models or professional judgment) in the assessment of risk of future harm to children (although some commentators have noted that developing a valid tool may be more straightforward than successfully implementing one) (Gambrill and Schlonsky, 2000; Barlow et al., 2012; D'Andrade et al., 2008).

22 At the same time, evidence reviews consistently find limited accuracy in predicting child maltreatment from the small number of tools that have been developed with the aim of predicting child maltreatment or ensuring that preventive early intervention services are targeted to families most in need (Leventhal, 1988; Peters and Barlow, 2003; Browne and Chou, undated; MacMillan, 2000; Nygren et al., 2004).

23 Responses to this finding vary across the reviews.

- The earliest review (Leventhal, 1988) concluded that prediction is feasible, but that improvements in the approach to the assessment of risk are needed, including further development and use of standardised assessment.

- The Canadian Task Force on Preventative Health Care concludes that there is sufficient evidence to recommend against screening for child maltreatment because of the high false positive rates and the harm associated with labelling parents as potential child abusers (MacMillan, 2000).

---

[4] Under the terms of the Privacy Act, the Privacy Commissioner must retain the capacity to conduct an independent review in the event of a complaint. As a consequence the Privacy Commissioner is not able to approve proposals such as this in advance. However, the Privacy Commissioner is able to signal any practices that are not permitted under the Act or that might pose a problem of perceived privacy risks.

- The United States Preventive Services Taskforce concludes that there is insufficient evidence to recommend *either for or against* screening (United States Preventive Services Taskforce, 2004) and that further research is needed (Nygren et al., 2004).

- One review concludes that the stigma attached to the use of screening instruments and difficulties with their application within a clinical context suggest the need for an alternative approach to the identification of parents in need of higher levels of intervention (Peters and Barlow, 2003).

- Another concludes that while none of the existing tools are precise enough to be used as screening instruments for child maltreatment, and labelling families as high risk would be unethical and stigmatising, screening can nevertheless be used to identify children and families in need of more support. This is provided the classification of families is expressed positively ("high priority for services" rather than "high risk" – a suggestion that could be taken up in implementation), and the interventions offered are well designed to address the problems identified (Browne and Chou, undated).

24 A 2013 United States Preventive Services Taskforce evidence update highlights a promising trial that involves screening for risk of abuse and neglect via a parent self-administered questionnaire (Nelson et al., 2013). The trial, which it rated as being of "fair" quality, includes risk assessment during the course of usual primary health care visits, training for physicians, and information resources and social work services for families (Dubowitz et al. 2009; Dubowitz et al. 2012). High-risk families in the intervention group had fewer care and protection reports and fewer episodes of severe or very severe physical assault than those in the usual care group three years after the intervention (Dubowitz et al. 2009).

25 The 2013 update (Nelson et al., 2013) and associated United States Preventive Services Taskforce recommendation statement (Moyer, 2013) reiterate the need for more research on effective methods for practitioners to identify children at risk or currently experiencing abuse or neglect, on effective interventions, and on whether there are unintended harms from screening, risk assessment and interventions.

26 To date, there has been very limited application of computational tools to support decision making in child protection and youth services (Schwartz et al., 2008). The potential for these tools to assist in predicting the risk of severe harm for children already reported for suspected abuse (Schwartz et al., 2004) and in prediction of youth re-offending (Schwartz et al., 2008) has been examined, and demonstrates considerable scope for improvement on existing risk assessment tools.

27 In contrast to standardised assessment tools where a checklist or questionnaire is applied to obtain a risk score, PRM tools generate a risk score based on existing administrative data.

28 Advantages of PRM tools are that they are developed using actual data for the specific population at risk and outcome of concern, and allow population-wide screening that is reasonably cost-effective. In contrast, standardised assessment tools tend not to be validated on the population to which they are applied (Schwartz et al., 2008), and can be under-utilised where they are resource intensive for front-line agents to administer, or not used as intended (Vaithianathan *et al*., 2012).

29 Within the social services, use of PRM tools based on large-scale administrative data is most advanced in healthcare (Panattoni et al., 2011). While utilisation of existing administrative data for early identification at the population level has been explored and advocated for (eg. Wu et al, 2004; Putnam-Hornstein and Needell, 2011), to our knowledge, the study by Vaithianathan et al. (2012) and this feasibility study are the first to

examine the potential for PRM tools that draw on existing administrative data to be used for early identification to support child maltreatment prevention efforts.

3. A base model that would apply around the time of birth

30 PRM tools use relationships in existing administrative data to estimate the likelihood of a future event. They generate an algorithm (a list of predictor variables and accompanying estimated weights) that can be applied to obtain a score which represents the probability of the event occurring. Scores in different bands can then be used to allocate cases to different levels and types of service response.

31 A base model was developed to begin the assessment of the feasibility and predictive performance of a population-wide PRM tool for identifying new-born children who are at high risk of maltreatment.

32 This section describes the linked research data, ethnicity information used in the study, the study population, outcome targeted, predictor variables and modelling strategy for the base model. Section 4 then assesses the base model's predictive performance, and compares this with other tools for early identification reviewed in the international literature.

Linked research data

33 In New Zealand there is no single unique identifier allocated to an individual against which all that individual's contact with different government administrative systems is recorded. In order to link data for the same individual across the different systems to support PRM, it would be necessary to link data using identity details such as names and dates of birth and other potentially identifying variables such as caregivers' and parents' names and dates of birth.

34 The research data assembled for the feasibility study linked, for children born from mid-2004 to mid-2012:

- birth notification[5] and registration information for the child and her or his parents and death registration information for the child held by the Registrar-General of Births, Deaths and Marriages

- benefit data from 1993 onwards for the child, for other children in the family, and for their parents or caregivers held by the Ministry of Social Development (MSD)

- Child Youth and Family (CYF) care and protection data from the early 1990s onwards for the child, for other children in the family, and for their parents or caregivers (relating to their own childhood) held by MSD

- data on sentences served by the child's parents or caregivers administered by the Department of Corrections

- a range of Ministry of Health data for the child and birth mother where they had a maternity record.

Ethnicity data

35 For the purposes of describing the predictive power and proportionality of the base model by the ethnic group of the child, birth registration information is used as this is considered

---

[5] Birth notification information includes birth weight and gestation.

the best available data source. Birth registrations provide all the ethnic groups to which the child belongs, as reported by the parents registering the birth.

36 In contrast, benefit data provide no information on the ethnic groups to which the child belongs and while they provide self-reported ethnic groups of the parents or caregivers, these data are not an ideal source for imputing the ethnicity of the child as in most cases a child included in benefit appears with a single parent or caregiver and no information is available on the ethnicity of the other parent. CYF data include the ethnicity of the child, but this information is available for only a minority of children, and may be reported by an individual making a notification of concern who does not have accurate or full knowledge of the child's ethnicities.

37 Where ethnic group is reported for models that do not draw on birth registration data, this is also sourced from birth registration information that can be linked to the child of interest.

Study population

38 The birth cohorts that are the focus of the modelling are children born over the four years from 2007 to 2010 inclusive. For these birth cohorts, the linked data provide at least two years of data on care and protection outcomes (covering the period up to the end of 2012). They also provide at least two and a half years prior to the child's birth over which to examine the care and protection histories of older siblings identified through birth records (because the linked research data cover birth registrations for children born from mid-2004).

39 Where the University of Auckland research examined the subset of children whose caregivers receive main benefits, here the study population is all children in a birth cohort who, by three months of age, had their birth registered or were included in a caregiver's benefit. This covers an estimated 94 percent of all New Zealand live-born children, and yields records for in excess of 60,000 children in each annual birth cohort.

40 For base models, the modelling was carried out on samples drawn from each birth cohort (these are described below).

Outcome targeted

41 Base models predict the risk of at least one substantiated finding of maltreatment in CYF data (emotional, physical or sexual abuse or neglect) by age two for children in the study population.

42 The outcome window (the period of time between the prediction being made and the outcome occurring) is short. This anticipates the need for any working model to prioritise the recency of data informing the model over a longer outcome window so as to ensure that the model, as much as possible, reflects data being generated by current administrative systems. It also focuses the predictions on an age group at which the rate of maltreatment-related death and hospitalisation (Gilbert et al., 2011; Child Youth and Family, 2006) and first instance of substantiated maltreatment (Wulczyn, 2009) is highest.

43 When CYF receives a notification or report of concern about a child, an assessment is made to determine whether further action is required. Further action can include:

- partnered response (which involves referring the family to a community-based service that can help)

- an investigation (where there is an allegation of serious child abuse or neglect which requires joint Police and CYF involvement)

- a Child and Family Assessment to identify the family's needs and develop a plan for how to respond (when there are safety, care or wellbeing concerns for a child or young person but where a forensically focused investigation is not required).

44 If the child is under 5 years old, the completion of a formal investigation or Child and Family Assessment is prioritised (Ministry of Social Development, 2012). Substantiated findings of abuse or neglect can be recorded by a social worker following either an investigation or a Child and Family Assessment.

45 The proportion of the study population with substantiated findings of maltreatment by age two increased from 2.3 percent for the 2007 cohort to 2.8 percent for the 2010 cohort based on conservatively linked data. The growth was the tail-end of an increase that began in the early 2000s. An important contributor to that increase was a change in Police procedures which resulted in a notification to care and protection services being made in respect of all children present at family violence incidents attended by the Police. This, together with recognition of exposure to family violence as a form of psychological abuse under the Domestic Violence Act (1995), was associated with a rise in the number and proportion of children with substantiated findings of emotional abuse (Mansell et al., 2011; CSRE, 2012).

46 A further change in procedures for handling Police family violence referrals was introduced in July 2010. From that date, Police family violence referrals that require no further action do not result in care and protection notifications, but instead are recorded as "contact records". This appears to have been associated with a levelling off of numbers of findings of emotional abuse (Ministry of Social Development, 2012).

Predictor variables – markers of risk and protection

47 The predictor variables considered for the base models are markers of risk and protection derived from data drawn from birth registration, benefit, care and protection and sentencing administrative systems. They are listed in Box 1 and described more fully in the technical companion paper.[6]

48 Selection of variables was informed by the literature on risk and protective factors for abuse and neglect (Runyan et al, 2002; Child Youth and Family, 2006; Kerslake Hendricks and Stevens, 2012; New Zealand Government, 2012), and by previous studies that examine associations between administrative data available at birth and subsequent reporting or substantiation of maltreatment (Wu et al., 2004; Putnam-Hornstein and Needell, 2011).[7]

49 Preliminary analysis established that ethnic group does not need to be included as a predictor in order to support PRM, adding only marginally predictive performance.

50 Where variables consider the history of parents, caregivers or other children, these look back only over the previous five years. This is in an effort to acknowledge that some risks

---

[6] These are variables considered *before* stepwise exclusion.

[7] Age of the mother at birth of the first child was considered, but we were not able to construct a satisfactory variable given the data available to us. For the majority of the cohort, we are limited to birth registration data which does not provide a basis for estimating age at first birth. Parity information recorded on the birth registration is restricted to a listing of "other children born from the same parent relationship (This means other children with the same mother AND father as the new child…"). In addition, the data purchased for the study only reach back to mid-2004, limiting our ability to comprehensively identify through the data themselves earlier births to the same mother.

might resolve over time. The exceptions to this rule are the parents' or caregivers' own care and protection history in childhood and adolescence, and whether there were recorded findings of behavioural or relationship difficulties for them in childhood or adolescence. This information is included regardless of its age. In practice, because events prior to the early 1990's do not tend to be captured by the available electronic data, these care and protection history variables are only available for younger parents.

Box 1 Variables considered in the base model

Child

Gender of child (male / female)

Low birth weight or pre-term (yes / no or unknown)

Other children of the parents and/or caregivers

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)

Other children with care and protection history in the last 5 years (yes / no)[8]

Family violence (no events / notification or contact record in 1 of the last 12 months / notification or contact record 2+ of the last 12 months)

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+)[9]

Care and protection history as a child (yes / no)

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%)

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)[10]

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)[11]

Corrections history in the last 5 years (no history / non-custodial sentence only / custodial sentence for non-violent crimes / custodial sentence for violent crimes)[12]

Community / office

CYF site or service centre (43 categories)

---

51 Information on whether other children have been the subject of Police family violence notifications or contact records in the previous 12 months and the intensity of that activity is included as a marker for exposure to family violence. Given the data linkage formed for the research, this information was only able to be consistently incorporated for children already known to care and protection services. In any implementation, it would be possible to establish systems to incorporate this information for all children, so long as these records continue to be generated and held.

---

[8] Excludes other children with Police family violence notifications or contact records but no other history.

[9] The small number of cases with missing information are included in the 30-34 category.

[10] Based on incapacities recorded when parents or caregivers have claimed Sickness or Invalid's Benefits. Substance abuse is prioritised so that this category applies where a person is recorded as having both a substance abuse and another mental health disorder.

[11] Based on benefit data.

[12] Sentences for violent crime are prioritised so that this category applies where a person is recorded as having both a custodial sentence for a violent crime and a custodial sentence for a non-violent crime.

52 The length of time the mother or primary benefit recipient with care of the subject child was supported by benefit in the previous five years (including time supported by benefits in adolescence for younger mothers and caregivers) was included as a marker for poverty and its persistence.[13] This risk marker would be expected to have associations with sources of parenting stress such as deprivation, household over-crowding[14] and poverty-related child illness and hospitalisation (Trenholme et al., 2012), and has been found to have associations with other parental risk factors, including adversity in their own childhood, low educational attainment, conduct disorder, and mental health and substance abuse disorders (Welch and Wilson, 2010).

53 A variable that indicates which CYF site the child would be served by given their place of residence is included in an effort to account for a number of dimensions on which local conditions might impact on outcomes for a child. These include, for example:

- the level of deprivation in the community and locality-specific relationships between alcohol outlet density and alcohol-related social harms (Cameron et al., 2013)

- social norms, the availability and co-ordination of services, and the degree to which there is a sense of collective responsibility for positive child development in the community (Daro and Dodge, 2009)

- organisational and case worker factors that can vary across local care and protection service offices, an important source of variation in decision making and recorded substantiations of maltreatment (Baumann et al. 2011).

54 "Capture" of risk is incomplete in some cases. In the base model, we are able to include markers indicating whether benefit data show a parent or caregiver has a history of mental health or substance abuse disorder, for example, but this will not capture all cases where there is a history of these disorders. Similarly, in the base model, address change information is only able to be included where the parents or caregivers have received welfare benefits. (The impact of broadening the capture of risk information using Ministry of Health markers of risk, including mental health and address changes, is explored in section 9 below.)

55 In some cases, the predictor variables are a direct measure of a known risk or protective factor (eg low birth weight or young maternal age). In other cases, factors that are known to be associated with maltreatment (such as parenting skills) are not able to be directly measured, but may be associated with markers available from administrative data.

56 Coverage of known protective factors is particularly weak, reflecting the limitations of the available data. It is not possible, for example, to include markers that relate to extended family or whānau support. With further data linkage, incorporation of information on the availability and use of services that have the potential to be protective (early childhood education or kōhanga reo, teen parent education units, parenting programmes and home visitation services for example) could be explored, but was not feasible in the current study.

---

[13] Benefit type was considered in preliminary modelling but was excluded from the final models because the predictive utility of this information will be limited until sufficient data accumulate under the reformed benefit type categories introduced in July 2013. A risk marker indicating whether or not the child is currently supported by benefit may continue to have predictive utility, however, and could be considered if PRM is trialled.

[14] The feasibility and predictive utility of incorporating administrative measures of household crowding could be explored if PRM is developed further.

57 Given incomplete capture and coverage of known risk and protective factors, the modelling is not able to identify factors that cause substantiated findings of maltreatment to occur, only risk and protective markers available from administrative data that have utility in prediction.

Modelling strategy

58 After testing a range of modelling techniques, stepwise logistic regression was used to generate the predictive algorithms. This modelling approach is transparent, and allows knowledge to be gained about the most useful predictors, a particular area of focus for the feasibility study.

59 A range of alternative modelling approaches were found to have as good predictive performance, and could be chosen for any future implementation.

60 In response to peer review, model testing included an examination of the impact of applying an alternative "multilevel" modelling approach to better account for the impact of local conditions. Results suggest that multilevel models provide a small improvement in predictive performance (although improvements were not consistently observed). Offsetting the gains in performance was an increase in the disproportionate representation of some sub groups compared to the stepwise logistic regression models.

61 The modelling was generally carried out on samples drawn from each birth cohort. At the population level, we are seeking to model the occurrence of events that are rare. A common approach for dealing with prediction of rare events and associated class imbalance in the source data is to change class distribution and create more balanced data for modelling.[15] In the base models we under-sample the majority class (children with no maltreatment findings by age two), while capturing all cases with findings of substantiated maltreatment.

62 Models were developed or "trained" on 70 percent of the records in the cohort sample and tested (or "validated") on the remaining, separate, 30 percent of records. This is a standard approach in PRM (Witten et al., 2011; Williams, 2011). To assess predictive accuracy by age five, the models developed were used to risk score all children born in 2007 in the study population.

63 Tests of sensitivity to the sampling and "training/validation" partitioning methods applied were undertaken. These showed little impact on predictive performance.

64 The technical companion report describes in more detail the results of model selection and sensitivity testing.

---

[15] Studies have shown that a more balanced dataset provides improved classification performance as compared with an imbalanced dataset and under sampling and over-sampling are effective methods for dealing with problems associated with class imbalance (Choi, 2010).

4. Predictive performance

65 Assessment of the true predictive accuracy of the models developed is not possible because much of the abuse and neglect that occurs goes undetected.

66 Table 1 summarises the predictive performance of the base model across different cohorts in terms of ability to predict that subset of maltreatment that is known to occur – that which results in substantiated findings of maltreatment.

67 In terms of area under the Receiver Operating Characteristics curve (AUR), the models perform well in ranking children from high to low risk when the predictive algorithm for each cohort is applied to the corresponding validation sample (Table 1).

68 The misclassification rate at three percent reported in Table 1 tells us that errors in prediction are made in around four percent of cases across the whole of the 2007 study population when children in the most at-risk three percent of that population are predicted to have a finding of maltreatment by age two.[16]

69 Cumulative lift figures show that if we used the models to identify the most at-risk three percent of the 2007 study population, we would find eight to nine times more children with findings of maltreatment by age two than if we randomly selected three percent of children.

70 The results in Table 1 assume the top three percent of cases represents the threshold for indicating high risk of later findings of maltreatment. Three percent approximates the proportion with actual findings of maltreatment by age two – the outcome targeted.

71 In practice, the threshold for deciding which children to refer for preventive services could be set at a lower or higher threshold than three percent (or could be set as a band if the decision was to exclude or offer some other form of service to those at the very highest risk). Table 2 presents additional results for a range of thresholds. It uses the model developed for the 2010 cohort to score the 2007 cohort, and compares the predicted risk scoring with actual outcomes up to age five.

72 An ideal prediction tool would identify every child who would subsequently go on to experience the outcome of concern (in this case, a substantiated finding of maltreatment) as high priority (100 percent sensitivity) and every child who would not have experienced the outcome of concern as low priority (100 percent specificity). In practice, no predictive test can achieve this and there is a trade-off between sensitivity and specificity when deciding on the threshold (Leventhal, 1988).

73 For illustrative purposes, had the five percent of children in the 2007 study population with the highest scores been defined as high priority children for referral:

- 30.5 percent of the children referred would have a substantiated finding of maltreatment by age five (positive predicted value)

- 31.6 percent of all the children who went on to have findings of maltreatment by age five would have been referred (sensitivity)

---

[16] And all other children are predicted not to have a finding of maltreatment. Misclassification rates are obtained by risk scoring the entire study population in the relevant birth cohort rather than the validation sample because the weighting of the validation sample to under-sample the majority class makes misclassification rates at thresholds other than 50% difficult to obtain using SAS Enterprise Miner.

- the referred group would account for 36.0 percent of *all maltreatment findings*, which is higher than the 31.6 percent of *children with findings* (counting events rather than children)

- in 96.4 percent of cases where a child had no findings of maltreatment by age five, they would not have been referred (specificity)

- 2,163 of the 3,114 children referred (69 percent) would have no substantiated finding of maltreatment (false positives) by age five (Table 2).

Table 1. Base models for birth cohorts 2007-2010
Stepwise logistic regression model results for children with birth registration or on benefit by 3 months of age[1]

	Model for cohort born in:			
	2007	2008	2009	2010
Number in cohort	62,273	62,900	62,639	63,176
% with findings of maltreatment by age 2[1]	2.3	2.5	2.6	2.8
Sample size for modelling	7,170	7,865	8,260	8,750
AUR (validation sample)[2]	87%	87%	89%	89%
(95% confidence interval)	(85%-89%)	(86%-89%)	(87%-90%)	(88%-91%)
Misclassification rate at 3% (scored study population)[3]	4.2%	4.1%	4.2%	4.2%
Cumulative Lift at 3% (scored study population)[4]	8.2	9.1	8.3	8.3

Notes:
(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.

(2) Area under the Receiver Operating Characteristic curve (AUR) is a diagnostic tool for evaluating the ability of a model to rank positive instances relative to negative instances. An AUR of 50% represents a worthless prediction that performs no better than the toss of a coin. The Receiver Operating Characteristics curve plots the proportion of actual positive outcomes that are predicted positives (the true positive rate or sensitivity) against the proportion of actual negative outcomes that are false positive predictions (the false positive rate or one minus specificity) at different thresholds.

(3) Misclassification rate at 3% is the proportion of *all* cases in which an error in prediction is made (either a false positive or a false negative prediction) using the top 3% of scores to define predicted positives. To allow comparison, the scored study population is the cohort born in 2007 across all models.

(4) Cumulative Lift is a measure of model performance that is useful when looking at the most at-risk part of the population. It gives the ratio of the sensitivity of the model at the given threshold to the sensitivity resulting from a random selection of individuals from the population. To allow comparison, the scored study population is the cohort born in 2007 across all models.

Predictive accuracy for Māori children

74 Table 3 replicates Table 2 for children in the cohort who are known to be Māori, based on birth registration ethnicity information supplied by the parents. Setting a threshold that referred children with the highest five percent of risk scores overall would see 11 percent of children known to be Māori referred. At this threshold:

- 33.2 percent of the Māori children referred would have a substantiated finding of maltreatment by age five (positive predicted value)

- 36.3 percent of all the Māori children who went on to have findings of maltreatment by age five would have been referred (sensitivity)

- in 91.8 percent of cases where a Māori child had no findings of maltreatment by age five, they would not have been referred (specificity)

- 1,271 of the 1,902 Māori children referred (67 percent) would have no substantiated finding of maltreatment (false positives) by age five (Table 3).

Table 2. Accuracy in identifying children with any findings of maltreatment by age five using different thresholds to define high risk
2010 models applied to children born in 2007 with birth registration or on benefit by 3 months of age[1]

Share of scores used to define high priority	Number of children referred	False positive[2] by age 5 (count)	False negative[3] by age 5 (count)	% referred children with findings by age 5 (PPV)[4]	% not referred children with no findings by age 5 (NPV)[5]	% all children with findings by age 5 referred (sensitivity)[6]	% all children with no findings by age 5 not referred (specificity)[7]
Top 1%	623	360	2,745	42.2	95.5	8.7	99.4
Top 2%	1,245	785	2,547	37.0	95.8	15.3	98.7
Top 3%	1,868	1,234	2,374	33.9	96.1	21.1	97.9
Top 4%	2,491	1,682	2,199	32.5	96.3	26.9	97.2
Top 5%	3,114	2,163	2,057	30.5	96.5	31.6	96.4
Top 6%	3,736	2,639	1,911	29.4	96.7	36.5	95.5
Top 7%	4,359	3,128	1,777	28.2	96.9	40.9	94.7
Top 8%	4,982	3,622	1,648	27.3	97.1	45.2	93.9
Top 9%	5,605	4,133	1,537	26.2	97.3	48.9	93.0
Top 10%	6,227	4,642	1,423	25.5	97.5	52.7	92.2
Top 11%	6,850	5,172	1,330	24.5	97.6	55.8	91.3
Top 12%	7,473	5,701	1,236	23.7	97.7	58.9	90.4
Top 13%	8,095	6,233	1,146	23.0	97.9	61.9	89.5
Top 14%	8,718	6,786	1,076	22.2	98.0	64.2	88.5
Top 15%	9,341	7,350	1,017	21.3	98.1	66.2	87.6
Top 16%	9,964	7,910	954	20.6	98.2	68.3	86.7
Top 17%	10,586	8,485	907	19.8	98.2	69.8	85.7
Top 18%	11,209	9,054	853	19.2	98.3	71.6	84.7
Top 19%	11,832	9,614	790	18.7	98.4	73.7	83.8
Top 20%	12,455	10,197	750	18.1	98.5	75.1	82.8
All	62,273	59,265	-	4.8			

Notes: See below.

Table 3. Accuracy in identifying Māori children with any findings of maltreatment by age five using different thresholds to define high risk
2010 models applied to children born in 2007 with birth registration or on benefit by 3 months of age[1]

Share of scores used to define high priority	Number of children referred	False positive[2] by age 5 (count)	False negative[3] by age 5 (count)	% referred children with findings by age 5 (PPV)[4]	% not referred children with no findings by age 5 (NPV)[5]	% all children with findings by age 5 referred (sensitivity)[6]	% all children with no findings by age 5 not referred (specificity)[7]
Top 1%	173	89	1,654	48.6	90.3	4.8	99.4
Top 2%	346	186	1,578	46.3	90.7	9.2	98.8
Top 3%	519	302	1,521	41.8	90.9	12.5	98.1
Top 4%	692	415	1,462	39.9	91.2	15.9	97.3
Top 5%	864	530	1,404	38.6	91.5	19.2	96.6
Top 6%	1,037	648	1,349	37.5	91.7	22.4	95.8
Top 7%	1,210	763	1,291	36.9	92.0	25.7	95.1
Top 8%	1,383	890	1,245	35.6	92.2	28.4	94.3
Top 9%	1,556	1,007	1,189	35.3	92.4	31.6	93.5
Top 10%	1,729	1,127	1,136	34.8	92.7	34.6	92.8
Top 11%	1,902	1,271	1,107	33.2	92.8	36.3	91.8
Top 12%	2,075	1,399	1,062	32.6	93.0	38.9	91.0
Top 13%	2,247	1,534	1,025	31.7	93.2	41.0	90.1
Top 14%	2,420	1,655	973	31.6	93.5	44.0	89.4
Top 15%	2,593	1,794	939	30.8	93.6	46.0	88.5
Top 16%	2,766	1,927	899	30.3	93.8	48.3	87.6
Top 17%	2,939	2,058	857	30.0	94.0	50.7	86.8
Top 18%	3,112	2,205	831	29.1	94.1	52.2	85.8
Top 19%	3,285	2,348	801	28.5	94.3	53.9	84.9
Top 20%	3,458	2,491	771	28.0	94.4	55.6	84.0
All	17,288	15,550	-	10.1			

Notes: See over.

Notes to Table 2 and 3

(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.

The following contingency table for the relationship between predictions and true outcomes was used to calculate measures of predictive accuracy:

Identified as high priority by PRM	Findings of maltreatment	No findings of maltreatment	Total
Yes	a	b	a + b
No	c	d	c + d
Total	a + c	b + d	n

(2) False positive (FP) count is the number of children in the predicted high priority group referred who subsequently have no substantiated findings of maltreatment. FP = b.

(3) False negative (FN) count is the number of children not in the predicted high priority group referred who subsequently have substantiated findings of maltreatment. FN = c.

(4) Positive Predictive Value (PPV) is the proportion of children in the predicted high priority group referred who subsequently have substantiated findings of maltreatment. PPV = a / (a + b).

(5) Negative Predictive Value (NPV) is the proportion of children in the predicted low priority group not referred who do not subsequently have substantiated findings of maltreatment. NPV = d / (c + d).

(6) Sensitivity measures the proportion of children who go on to have substantiated findings of maltreatment who are correctly identified as being high priority and referred. Sensitivity = a / (a + c).

(7) Specificity measures the proportion of children who do not go on to have substantiated findings of maltreatment who are correctly identified as being low priority and not referred. Specificity = d / (b + d).

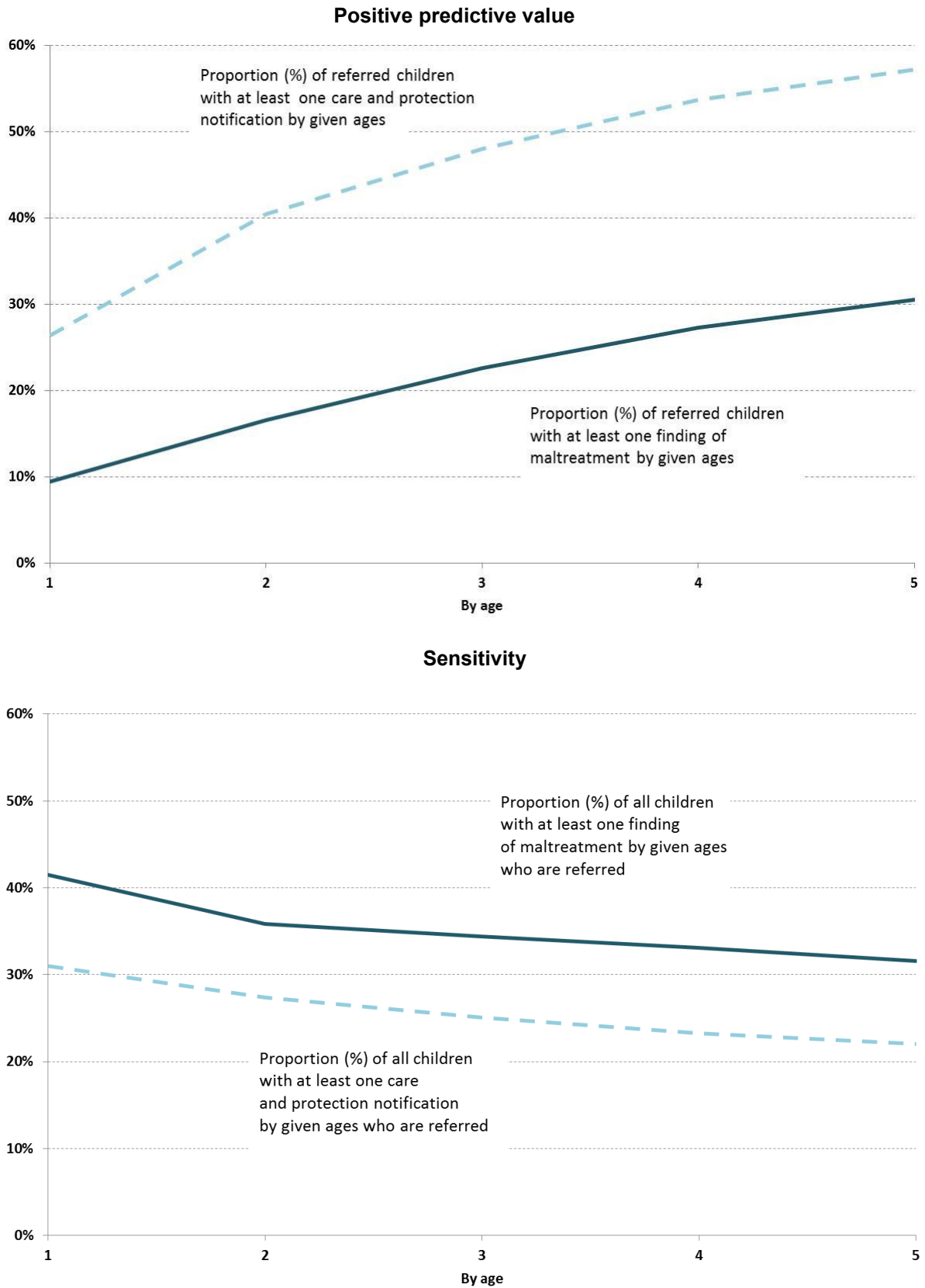*Positive predictive value and sensitivity with increasing age*

75 Figure 1 shows that positive predictive value builds with increasing age for the five percent of children in the 2007 study population with the highest scores. Projecting forwards from current trends, by age 10 the proportion of children referred using this as the threshold who would have subsequent findings of maltreatment might reach four in ten.

76 Sensitivity declines with age. Referring the five percent of 2007-born children with the highest scores would have captured 41.5 percent of children with substantiated findings of maltreatment by age one. This falls to 31.6 percent of children with substantiated findings by age five and projecting forwards might be around one in five by age 10.

77 Substantiated and unsubstantiated notifications may not be dissimilar in terms of the presence of risk factors or in the rate of re-notification (Drake, 1996; Hussey et al., 2005; Fluke, 2009). Notifications that do not result in substantiated findings may involve high levels of harm, but insufficient evidence, moderate levels of harm that do not meet the threshold for findings to be recorded, or events where no harm has occurred but that suggest risk of future harm, in addition to situations where the child is at low risk of harm.

78 With this in mind, it is useful to also consider accuracy in predicting which children would be notified. Had the five percent of children in the 2007 study population with the highest scores been defined as high priority children for referral, by age five, 57.2 percent of these children would have at least one notification (Figure 1). Projecting forwards, the proportion who are notified might rise to seven in ten by age ten. The referred group would account for 22.0 percent of *all notified children* by age five (Figure 2) but 31.4 percent of *all notifications* by that age.

Figure 1. Positive predictive value and sensitivity, where the five percent of children with the highest scores are high priority for referral

2010 models applied to children born in 2007 with birth registration or on benefit by 3 months of age[1]

Positive predictive value

Proportion (%) of referred children with at least one care and protection notification by given ages

Proportion (%) of referred children with at least one finding of maltreatment by given ages

By age

Sensitivity

Proportion (%) of all children with at least one finding of maltreatment by given ages who are referred

Proportion (%) of all children with at least one care and protection notification by given ages who are referred

By age

Notes:
(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.

79 A tool based on the circumstances of a child at birth has better sensitivity in infancy than in later years because it misses newly emerging circumstances that threaten or protect the child as he or she ages (Leventhal, 1988). If PRM is to be implemented, the feasibility of additional screening based on information that becomes available at older ages (including, for example, caregiver change data available from the benefit system (Vaithianathan et al., 2012) or education data) would be explored.

80 Implementation of a PRM tool in the manner outlined in the White Paper would require a decision to be made about the threshold for referral. This decision is likely to be informed by a range of factors, including the capacity of the Children's Teams and the availability of services, as well as predictive accuracy.

81 Choice of threshold would need to recognize the trade-off between positive predictive value and sensitivity. Setting a threshold to maximize the proportion of the referred group who would go on to experience the outcome (at, say five percent) means identifying fewer of the total pool of children who would experience this outcome (i.e., a higher number and rate of false negatives). On the other hand, setting a lower threshold to identify more of the total pool of children who would go on to experience this outcome (at, say 10 percent) means that fewer of those referred would actually go on to experience this outcome (i.e., a higher number and rate of false positives).

82 The output from a PRM is a probability and represents the likelihood of an uncertain future event. Regardless of the threshold chosen, models will identify as low priority some future children who are the victims of high profile abuse or neglect, and identify as high priority children who would never experience abuse or neglect. This needs to be acknowledged as an inevitable limitation of any early identification approach.

Comparison with other tools for early identification

83 Leventhal (1988) provides a guide to what should be able to be achieved with a "good" predictor for child maltreatment, accepting that perfect prediction is impossible given that determinants of human behaviour are complex, and factors affecting risk change over time. He defines a good predictor as an assessment that has a positive predictive value of 25 percent or more, sensitivity of 40 to 60 percent, and specificity of 90 percent or more (Leventhal, 1988).

"A positive predictive accuracy of 25 per cent means that only one quarter of the children who are identified as being at high risk will subsequently be maltreated. However, another large group of the high-risk families (perhaps 25 per cent or more) may have parenting problems other than maltreatment and thus will likely benefit from some form of intervention. Initially, services could be offered to all who are considered to be at high-risk and then the intensity of the intervention could be adjusted to the family's needs and functioning over time. A sensitivity of 40 to 60 per cent means that only about half of the children who are subsequently maltreated will be included in the high-risk group. These rates of the positive predictive accuracy and sensitivity should result in a very high specificity." (Leventhal, 1988: 157).

84 Few tools developed previously meet all three of these criteria. In one review, only one of the tools reviewed met all three criteria and only two combined specificity over 80 percent with a positive predictive value above 25 percent (Peters and Barlow, 2003).

85 Assessed in terms of its ability to predict findings of maltreatment by age five, the base model presented here meets all three criteria where the threshold for defining children as high priority captures the seven to ten percent of children with the highest scores (Table 2).

86 Where the threshold captures the five to six percent of children with the highest scores, the base model meets the positive predictive value and specificity criteria by age five, and meets the sensitivity criterion for maltreatment findings to age one but falls short of meeting that criterion for maltreatment findings to age five.

5. Data linkage

87 An important question for this feasibility study is whether administrative data from across agencies and across system databases within agencies can be successfully linked in a timely and accurate manner to support PRM.

88 The base model uses data assembled using linkages that are "conservative" in that child data from two systems or two agencies are linked only if there is a very high level of agreement between the child's identity details in the two systems and the given name of at least one parent or caregiver in the two systems is the same. Parent or caregiver identities are only linked if there is a very high degree of agreement in their identity details (see the technical companion report for the rules applied).

89 The aim of the conservative linkage is to approximate an automated administrative approach which would seek to minimise the likelihood of erroneously linking one individual's information to another individual (false-positive linkages). Inevitably, this leads to some instances in which information that did in fact relate to the same individual fails to be linked (false-negative or "missed" linkages).

90 For comparison, a less conservative linkage of the data was also formed. This allowed links to be made where conservative linking criteria were not met but the chance of the data relating to the same individual was high. The preliminary research on PRM undertaken by Vaithianathan et al. (2012) was based on a less conservative linkage of data.

91 This section summarises the results from a clerical check of the accuracy of the linkages, and tests the sensitivity of the base model's predictive accuracy to the level of conservatism applied in the linkage of data.

A clerical check of accuracy in the linkage

92 To assess the quality of the linkages formed for the study, a clerical check was undertaken for a stratified random sample of linked records for 527 children born in 2010[17] and 176 of their caregivers.[18]

93 The 2010 cohort was chosen for the checking because it allowed stratification by recorded maltreatment outcomes by age two (so as to over-sample cases that are more likely to be scored as high risk and for whom predictions could be informed by large numbers of linkages). It is a reasonably recent cohort for which data capture and data quality would closely approximate that of the current administrative systems.

94 The checkers examined the quality of the identity linkages across the benefit, care and protection, Corrections and birth registration systems.[19] Where the data being linked included benefit or care and protection identities, searches were carried out by systems experts using the relevant administrative systems. Where it included other agencies' data, searches for possible missed and false positive matches were conducted within the identity data supplied.

---

[17] Records were sampled for children who appeared in each of the administrative systems included in the check within 3 months of birth with the aim of approximating the study population used for this research. See the Technical Appendix for details.

[18] This involved checking linkages for the parents and caregivers of around one in five of the children in each stratum.

[19] At the time the clerical check was performed the project team had not received the Ministry of Health data so identities from this system were not included in the checking reported here.

95   Table 4 shows that in the vast majority of cases, the linkages formed for the children agreed with those arrived at on clerical inspection, with a higher rate of agreement for the less conservative data linkage (99.6 percent) than for the conservative linkage (95.3 percent). All the errors found involved missed linkages.

Table 4. Proportion of cases in which child linkages agree with clerical check
Weighted estimates based on a stratified random sample of children born in 2010 (n = 527)

Group	Conservative linkage (%)	Less conservative linkage (%)
Across all children	95.3 (95% CI 93.1 - 97.0)	99.6 (95% CI 97.5 - 100.0)
Children with findings of maltreatment by age 2	82.4 (95% CI 77.3 - 86.2)	99.0 (95% CI 95.0 - 99.8)
Children with no findings of maltreatment by age 2	95.8 (95% CI 93.5 - 97.5)	99.6 (95% CI 97.4 - 100.0)

Notes:
95% CI = 95 perecent confidence interval for the estimates.

96   The area that posed the most problems was the linkage of care and protection records for a child with records from other systems, with agreement between the conservative linkage and the clerical check in only 82.4 percent of cases where the child had findings of maltreatment by age two. Virtually all of the unmatched cases could be successfully linked in the less conservative linkage, with agreement between that linkage and the clerical check in 99.6 percent of cases.

97   Difficulties in linking care and protection data reflect a general practice of establishing care and protection administrative identities without identity verification. Missed matches frequently occur where information about the child's identity omits or estimates key information (such as date of birth), is subject to some inaccuracy (for example, in the reporting of the child's name), or uses informal rather than formal names for the child, parents or caregivers.

98   In contrast, children included in a caregiver's benefit must have their identity verified against a birth certificate, either before or shortly after the granting of the child inclusion. As a result, the child identity information in the benefit system is of relatively good quality.[20]

99   When looking at the accuracy of the links for a child and all the adults associated with them, the rate of agreement with links arrived at on clerical inspection was lower at 81.9 percent for the conservative linkage and 93.8 percent for the less conservative linkage (Table 6). The lower accuracy reflects the increased likelihood that any one of the people involved could have had an error in the data linkage, and the higher rate at which errors were found for the adults reflecting their longer, and in some cases more complex, histories.

---

[20] However, it is apparent that birth certificates are not always referred to as a small number of children are included in benefit for months or years before their birth is registered. See Figure 4.

Table 5. Proportion of cases in which child and all associated adults' linkages agree with clerical check
Weighted estimates based on a stratified random sample of children born in 2010 (n = 96)

	Conservative linkage (%)	Less conservative linkage (%)
Across all children	81.9 (95% CI 67.7 - 91.7)	93.8 (95% CI 80.4 - 98.8)

Notes:
95% CI = 95 perecent confidence interval for the estimates.

100 For the adults, all of the errors found in the conservative linkage were missed matches. In the less conservative linkage, the error rate for adults was lower, but errors found included false positive linkages (4 for the 176 adults checked) where one individual's information was erroneously linked to another individual, as well as missed matches (2 for the 176 adults checked).

101 Depending on the number of other children in the family, the analysis data for one child could potentially draw on as many as 20 to 30 individual-to-individual linkages between systems. The estimates presented here (which only consider accuracy of linkages for the reference child and the related adults) should therefore be viewed as conservative estimates of the proportion of cases in which the information informing risk scoring might, under a conservative data linkage, be incomplete as a result of missed linkages.

102 While for the children no false positive matches were found in the clerical review, these were uncovered in other examination of the data. In some of these cases, information for twins was erroneously linked due to the similar patterning of the two children's names (for example, twins with the same given name, different middle names, and same surname were erroneously linked and treated as one child under both conservative and less conservative linkages).

Accuracy of predictions with a less conservative data linkage

103 Table 6 compares predictive accuracy and the composition of the referred group when the base model predictive algorithm (derived using the conservative data linkage for the 2010 cohort) is applied to the 2007 study population with data assembled using a conservative data linkage (column A) then applied to that same study population with data assembled using a less conservative linkage (column B).

104 Shifting from the conservative to the less conservative data linkage, the number of children risk scored falls by around 1,000 (reflecting an increase in cases where records are viewed as relating to the same rather than different children), and the proportion of children in the cohort overall with findings of maltreatment by age five increases from 4.8 percent to 6.0 percent (reflecting the greater success the less conservative data linkage has in linking care and protection records for the children).

105 Were the 3,000 children with the highest scores to be defined as high priority for referral (close to five percent of the conservatively linked population), positive predictive value for findings of maltreatment by age five increases from 30.7 to 37.2 percent comparing the two linkages. Sensitivity changes little, suggesting that missed information on findings of maltreatment under the conservative linkage occurs for children at both lower and higher assessed risk.

Table 6. Risk scoring the 2007 cohort using the 2010 base model with conservative and less conservative linkages of data

2010 models applied to children born in 2007 with birth registration or on benefit by 3 months of age

	A Risk scoring based on conservatively linked data	B Risk scoring based on less conservatively linked data
Number in 2007 cohort	62,273	61,282
% with findings of maltreatment by age 5	4.8	6.0
Using the 3,000 children with the top scores to define high-priority for referral:		
% referred children with findings by age 5 (PPV)	30.7	37.2
% all children with findings by age 5 referred (sensitivity)	30.7	30.5
% all children with no findings by age 5 not referred (specificity)	96.5	96.7
Profile of referred children (%):		
Other children with a care and protection history in the last 5 years	59.9	65.5
Single parent	87.7	88.5
Mother or caregiver aged under 25	54.6	54.3
Parents or caregivers with a care and protection history as a child	57.2	60.8
Child seen on benefit by 3 months of age	92.3	93.0

106 The largest changes in the composition of the 3,000 children with the highest scores comparing the two linkages are in the proportions with other children or parents or caregivers with a care and protection history. This is consistent with care and protection data being the area where information is most likely to be missed with a conservative data linkage.

107 Of all the children who were among the 3,000 with the highest scores in the conservative linkage:

- 41 (1.4 percent) were also among the 3,000 children with the highest scores in the less conservative linkage, but appeared twice among the 3,000 children with the highest scores in the conservative linkage because of a missed identify match

- 403 (13.4 percent) did not appear among the 3,000 children with the highest scores in the less conservative linkage (other children had been assessed as at higher risk).

108 These results highlight the potential for errors in prediction arising from the need to link data across administrative systems, and the potential for gains in predictive accuracy to be made if the level of conservatism in data linkage is able to be reduced.

109 Implementation of PRM would require careful decisions to be made about the degree of conservatism in data linkage. It needs to be acknowledged that an approach designed to minimise false positive linkages will inevitably mean that some information that should inform the assessment of risk for a child but is held under identities that are more difficult to link with certainty will be missed.

110 Given the potential for error, a system for review of the linkages that inform the risk scoring would be recommended before details of a high risk child were used as part of any working PRM. Such a system would aim to find cases where information has been missed and there is identity information in the source administrative systems that needs correction. It would also aim to prevent an approach being made to a family where a linkage error means an unrelated person's information has inflated the child's risk score. Having an effective system of review in place may allow a less conservative approach to data linkage to be taken, and improvements in predictive accuracy to be realised.

6. Proportionality of predictions

111 A concern with PRM is that it might be driven by (and, in turn, intensify) an over-representation of Māori and other social groups in the care and protection system that is disproportionate to their share of maltreatment – that it might have a "ratchet" effect (Harcourt, 2006) that feeds a cycle of bias in surveillance and, as a result of that surveillance, in findings of maltreatment. At the same time, it may feed a cycle of under-serving children at risk in other population sub-groups.

112 This section examines whether the representation of selected social groups among those who might be referred by a PRM is proportionate to their share of the population of children known to be maltreated, based on CYF administrative data on substantiated findings of abuse and neglect.

113 What is not addressed here but is being examined in further research is the question of whether the representation of these groups would be disproportionate relative to their real share of maltreatment, acknowledging that known maltreatment accounts for only part of the abuse and neglect that occurs (Munro, 2011; Gilbert et al., 2009; Runyan et al., 2002; Child Youth and Family, 2006). Whether the over-representation of some sub-groups of children among those with findings of maltreatment genuinely reflects heightened maltreatment risk, or reflects bias in surveillance and substantiation, is the subject of unfolding understanding (in the case of ethnic over-representation - Drake et al., 2011; Putnam-Hornstein, 2012) and continued uncertainty (in the case of over-representation of children and families with prior care and protection services - Fluke, et al., 2008).

Proportionality for Māori

114 Table 7 looks at children born in 2007 for children for whom ethnicity is known from birth registration information. It compares the ethnic composition of this cohort with that of the subgroup of children known to be maltreated by age two, and that of the 3,000 children with the highest risk scores when the cohort is risk scored using the 2010 base model. (Recall that ethnic group is not included in the model as a predictor.)

115 It is not currently known whether the high representation of Māori children among those known to be maltreated accurately reflects their real share of maltreatment. Against a historical back drop of colonisation, rapid urbanisation and concentration in industries most affected by job loss, Māori are disproportionately represented in deprived communities, and bear a disproportionate share of a range of social harms that are risk factors for child abuse and neglect (Cooper and Wharewera-Mika, 2009; Cram, 2012; Child Youth and Family, 2006). Research under way is assessing whether, in addition to the impact of these risks, there is any indication that bias in reporting or substantiation also contributes to their over-representation among children known to be maltreated.[21]

116 Of those whose ethnicity is known, Māori children would be slightly over-represented among the children referred, relative to their share of known maltreatment (comprising

---

[21]   Recent findings from United States research suggest that the higher rate of contact with the child welfare system for black children in that country largely reflects real differences in the underlying incidence of maltreatment, and this, in turn, reflects stark differences in the burden of poverty and associated social harms (Bartholet et al., 2011; Drake et al., 2011; Putnam-Hornstein et al., 2013b). At the same time, there is some evidence to suggest that for Canadian children subject to child maltreatment investigations, placement occurs more frequently where the investigating agency proportionately has a higher aboriginal caseload, potentially pointing to bias in child protection outcomes resulting from inequities in child welfare resourcing (Chabot et al., 2013; Fallon et al., 2013). For a collection of studies and discussion, see http://www.law.harvard.edu/programs/about/cap/cap-conferences/rd-conference/rd.conference.papers.html.

69.2 percent of the five percent of children with the highest risk scores compared with 60.9 percent of children known to be maltreated by age two). Children of Pacific, Asian and NZ European ethnicities would be under-represented. One way of addressing this imbalance is to deploy separate models for Māori children and children who are not known to be Māori and select the highest risk scored children for referral from each model in proportion to shares of known maltreatment. This approach is explored below.

Table 7. Ethnic composition of children overall, children with findings of maltreatment by age two, and the 3,000 children with the highest PRM scores

2010 models applied to children born in 2007 with birth registration or on benefit by 3 months of age[1]

Ethnic group(s) (total response)[2]	% of all children	% of children with findings by age 2	% of the 3,000 children with the highest PRM scores
Composition of children in group with non-missing ethnicity:			
Māori	28.4	60.9	69.2
Non-Māori	71.6	39.1	30.8
NZ European	71.4	53.2	49.9
Pacific	14.7	21.7	18.9
Asian	11.3	4.2	2.0
Other	1.8	0.8	1.0
% of all children in group with missing ethnicity data [3]	2.4	2.8	10.5

Notes:
(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.
(2) Ethnic groups are as recorded on the birth registration for the child.
(3) Ethnicity data are missing for children for whom there is no linked birth registration data, considering all birth registrations made by June 2012. Possible reasons for the apparent absence of a birth registration (and as a result missing ethnicity data) include (i) being not born in New Zealand (ii) being born in New Zealand and remaining without a birth registration at June 2012, and (iii) being born in New Zealand and with a registered birth but there being differences in identity information in the births and benefit data which meant that the records could not be linked with the data linkage algorithms applied.

Proportionality by other characteristics

117 On a range of other measures, PRM would refer children in numbers that do not reflect their share of known maltreatment (Table 8). It would, for example, refer too many new-born children:

- with parents or caregivers who receive benefit (comprising 92.3 percent of the five percent of children with the highest risk scores compared with 79.0 percent of children known to be maltreated by age two)

- in a family where other children have had contact with care and protection services (59.9 percent compared with 34.9 percent)

- with parents or caregivers with known childhood contact with care and protection services (57.2 percent compared with 43.2 percent)

118 The base model is best able to predict high risk where there is prior contact with social agencies or a sibling history of contact. Where there has been little or no such contact, the model relies only on the information that can be derived from birth registration and birth notification data (age of mother, for example). Consistent with other research, this information, on its own, is not as strong in distinguishing high from low risk (Leventhal, 1988), and children for whom only this information is available are under-represented among those with the highest predicted risk, relative to their share of known maltreatment.

Table 8. Profile of children overall, children with findings of maltreatment by age two, and the 3,000 children with the highest PRM scores
2010 models applied to children born in 2007 with birth registration or on benefit by 3 months of age[1]

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

Notes:
(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.
(2) If children are identified through benefit data first time caregiver status is assumed to apply if the caregiver has never previously received benefit with a dependent child, looking back to 1993. If children are identified through birth data, first time caregiver status is assumed to apply if the birth registration indicates that there have been no previous children from the same parent relationship and there are no New Zealand registered previous births to the same mother (looking back to 2004). Note that this will overstate cases where there are no other children where it misses previous children from a different parent relationship and where it misses previous New Zealand registered births that occurred prior to 2004.

7. Sensitivity to use of separate models to control disproportionality

119 In any implementation, over-representation relative to known maltreatment could be addressed by deploying separate models for population sub-groups and selecting numbers from each in proportion to their share of known maltreatment.

120 Table 9 presents measures of predictive accuracy when the 3,000 children in the 2007 cohort with the highest predicted risk are identified using this approach. It demonstrates the combination of separate models for:

- Māori new-borns versus new-borns who are not Māori or for whom ethnicity is not recoded

- new-borns in families where there are other children with a care and protection history versus those who were not

- new-borns supported by benefit by three months of age versus those who were not.

121 The technical companion provides details of the six separate models and their predictive performance when assessed using the validation sample for each sub-population and the scored study population (Appendix 2).

122 Compared with the base model, combining separate models has little impact on overall positive predictive value and sensitivity (Table 9).

Table 9. Combining separate models to identify the 3,000 children with the highest PRM scores
2010 models applied to children born in 2007 with birth registration or on benefit by 3 months of age[1]

	% referred children with findings by age 5 (PPV)	% not referred children with no findings by age 5 (NPV)	% all children with findings by age 5 referred (sensitivity)	% all children with no findings by age 5 not referred (specificity)
Base model	30.7	96.5	30.7	96.5
Combining separate models for:				
Māori child versus not [2]	30.1	96.4	30.0	96.5
With other children with care and protection history versus not	29.6	96.4	29.5	96.4
Child supported by benefit versus not	30.4	96.5	30.3	96.5

(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.
(2) Ethnic groups are as recorded on the birth registration for the child. Children can have more than one ethnic group recorded. Ethnicity data are missing for children for whom there is no linked birth registration data, considering all birth registrations made by June 2012. Where ethnicity is missing, the child is included in the "not Māori or ethnicity not recorded" model.

123 In some cases, addressing disproportionality along one dimension reduces disproportionality along others (for example, addressing disproportionality in the representation of new-borns supported by benefit reduces the disproportionate

representation of children of single parents) (Figure 2)[22]. In other cases, addressing disproportionality along one dimension intensifies that occurring on other dimensions (for example, addressing disproportionality in the representation of new-borns in families where there are other children with a care and protection history increases the disproportionate representation of children with parents or caregivers with a care and protection history in their own childhood and increases the disproportionate representation of children with younger parents).

124 The results show that addressing disproportionality on a particular dimension of concern, and in particular bringing the representation of Māori children and children in families where there are other children with a care and protection history into line with their share of known maltreatment, is possible to achieve without notable loss of predictive accuracy. At the same time, a solution that addresses disproportionality on all dimensions of concern is unlikely.

Need for further development

125 If there is a decision to take PRM to trial and to operationalize separate models to control disproportionality, further development would be required.
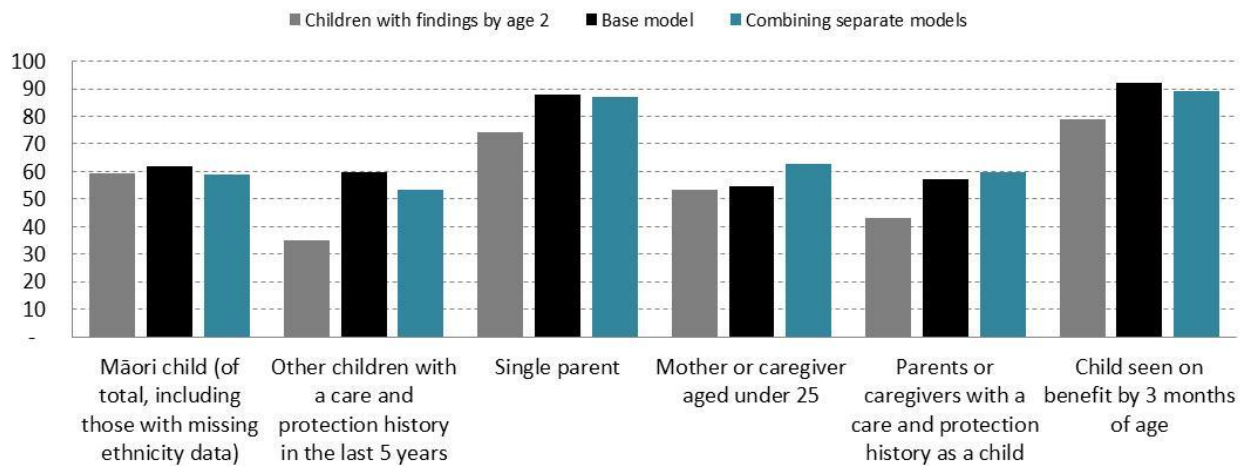
- In the demonstration of a separate model for Māori new-borns outlined above, ethnicity information is based on total response ethnicity data available from birth registrations to June 2012. This overstates the information on the ethnicity of the child that would be available to any working model (from Table 8, for 6.3 percent of children overall and 35.5 percent of the 3,000 children with the highest risk scores, the birth was not registered at three months of age). Options for operationalization include basing the separate modelling on the ethnicity of the parents or caregivers, or using the birth registration ethnicity information where it is available and imputing child ethnicity based on that of the parents or caregivers in remaining cases. If the decision is to proceed without birth registration data, the ethnicity of the child could not be used as the basis for separate models, and the ethnicity of the parents or caregivers would need to be used (noting that where benefit income is received, only one parent or caregiver is known in most cases).

- In the demonstration of separate models generally, the predictor variables and categories used in the modelling could be better tailored to the relevant sub-population. For example, in the model for new-borns in families where there are other children with a care and protection history, a finer breakdown of the nature of that history and tailoring of other variables could improve model performance.
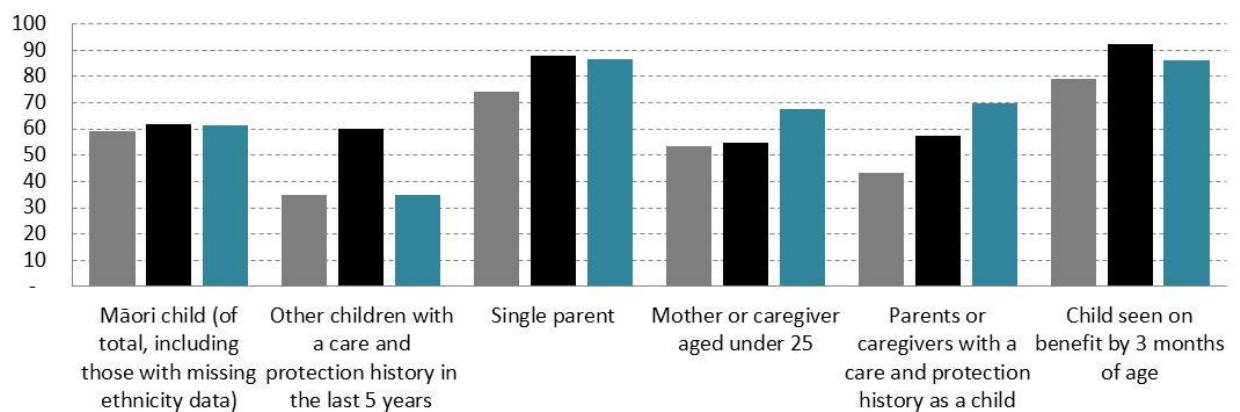
---

[22] Figure 2 presents proportionality for Maori children considering their representation in the total child population (including children for whom ethnicity is missing) where Table 7 considers their representation in the population with non-missing ethnicity information.

Figure 2. Proportionality combining separate models to identify the 3,000 children with the highest PRM scores

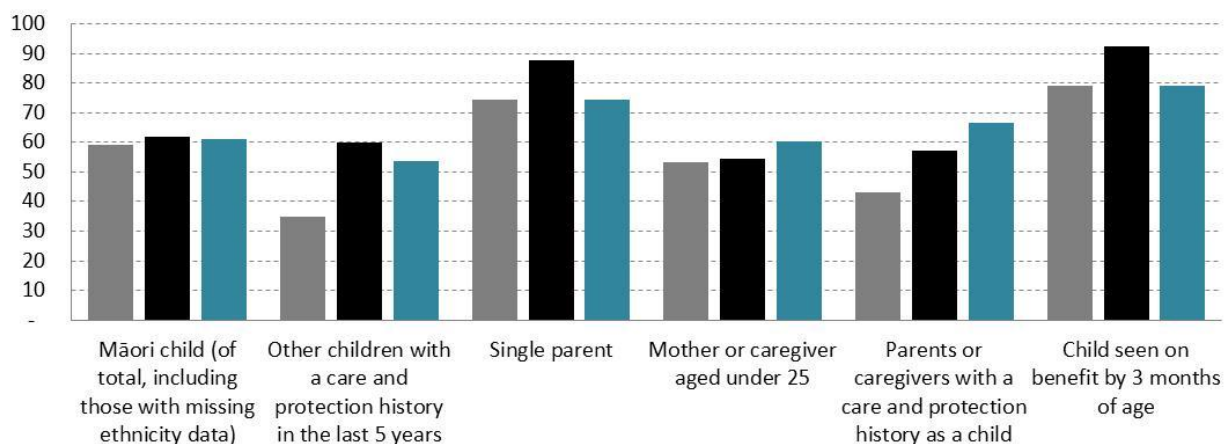2010 models applied to children born in 2007 with birth registration or on benefit by 3 months of age[1]

Combining separate models for Māori child versus not

Combining separate models for children where other children with a care and protection history versus not

Combining separate models for children supported by benefit versus not

(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.

8. Sensitivity to alternative specifications of the target variable

126 This section tests sensitivity to alternative specifications of the target variable, acknowledging that there are possible concerns with use of substantiated findings of maltreatment, the outcome targeted in the base model, as a proxy for the real occurrence of maltreatment.

- Some research suggests that reports or notifications of maltreatment, rather than substantiations, may provide a more accurate demarcation of whether harm has occurred (Drake, 1996; Hussey et al., 2005; Fluke, 2009).

- Past changes in administrative processes in New Zealand have greatly increased the incidence of findings of emotional abuse (CSRE, 2012) highlighting the role that administrative processes can play in shaping the rate of maltreatment substantiations.

127 Table 10 compares models for a range of target variables. The base model (A) and models targeting notifications (B), investigations or Child and Family Assessments (C) and findings of physical or sexual abuse or neglect (maltreatment excluding the emotional abuse category (D)) all have AURs sitting between 86 and 90 percent indicating good ranking ability. AUR is lower at 79 percent for the model targeting either physical abuse or hospitalisation for maltreatment or marker injuries.

128 The models excluding emotional abuse from the target (D) and targeting physical abuse or maltreatment or marker injury hospitalisations (E) have a lower misclassification rate and model D has high cumulative lift. These improvements partly reflect the mechanical effects of targeting outcomes that are more rare.

129 Using the 2010 models to risk score the cohort born in 2007 and identify the 3,000 children with the highest scores:

- positive predictive value for the outcome targeted varies in line with the rarity of the events

- sensitivity, in terms of capture of the children with the outcome targeted among those who would be referred, is highest for the model targeting physical or sexual abuse or neglect (D) and for the base model (A).

130 In addition to the base model, models targeting a range of care and protection outcomes perform well in terms of AUR and could be explored further in any operationalisation.

131 Models that predict which children will be notified to CYF or will be the subject of an investigation or Child and Family Assessment have high positive predictive value, and while they have relatively low sensitivity for the outcomes they target when assessed at a given threshold relative to the base model, sensitivity is higher at thresholds more in line with the population prevalence of these events.

132 A model that predicts which children have findings of maltreatment excluding the more volatile emotional abuse category inevitably has lower positive predictive value than the base model for the outcome it targets, but this builds with age, and otherwise the model performs well with the highest sensitivity of all the models at a given threshold.

133 Due to poorer model performance and the rarity of the events targeted, development of a model for predicting which children will have substantiated findings of physical abuse or will be hospitalised for maltreatment related injuries or marker injuries is not feasible, with very low positive predictive value and relatively low sensitivity.

Table 10. Base model compared with models with alternative target variables

Developed for children born in 2010[3] with a birth registration or seen on benefit by 3 months of age and used to risk score children born in 2007[1]

	A	B	C	D	E
	Base model (targeting substantiated maltreatment)	Targeting notifications [2]	Targeting investigations or Child and Family Assessments	Targeting substantiated physical or sexual abuse or neglect	Targeting substantiated physical abuse or maltreatment or marker injury hospitalisation [3]
Number in cohort	63,176	63,176	63,176	63,176	188,783
% with targeted outcome by age 2	2.8	8.0	5.5	1.1	0.1
Sample size for modelling	8,750	63,176	17,326	3,415	3,565
AUR (validation sample)	89%	86%	87%	90%	79%
(95% confidence interval)	(88%-91%)	(85%-87%)	(85%-89%)	(88%-92%)	(76%-83%)
Misclassification rate at selected threshold[4] (scored 2007 study population)	4.2%	9.3%	6.6%	1.8%	1.3%
Cumulative Lift at selected threshold[4] (scored 2007 study population)	8.3	5.1	6.5	12.1	7.9
% PPV for the targeted outcome by age 2 at selected threshold[4] (scored 2007 study population)	19.1	37.9	29.5	12.0	2.4
% sensitivity for the targeted outcome by age 2 at selected threshold[4] (scored 2007 study population)	24.9	41.0	32.3	12.1	7.9
% specificity for the targeted outcome by age 2 at selected threshold[4] (scored 2007 study population)	97.5	94.6	96.3	99.1	99.0
Using the 2010 outcome specific model to risk score the 2007 cohort and the 3,000 children with the top scores to define high-priority for referral:					
% referred children with targeted outcome by age 5 (PPV)	30.7	50.4	48.1	15.9	3.9
% all children with targeted outcome by age 5 referred (sensitivity)	30.7	18.7	25.4	34.8	26.9
% all children without targeted outcome by age 5 not referred (specificity)	96.5	97.3	97.2	95.9	95.3
% referred children with findings of maltreatment by age 5 (PPV)	30.7	31.7	30.7	30.9	28.7
% all children with findings of maltreatment by age 5 referred (sensitivity)	30.7	31.6	30.6	30.8	28.7
% all children without findings of maltreatment by age 5 not referred (specificity)	96.5	96.5	96.5	96.5	96.4

Notes:

(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.

(2) Including Police FV notifications, and contact records for children known to Child Youth and Family.

(3) Includes maltreatment injury hospitalisations by age 2 and hospitalisation for intracranial injury or long-bone fracture by age 1. Excludes short-stay emergency department hospital admissions. For this model, the rarity of the target variable made it necessary to pool data over cohorts in order to create the sample for modeling. The number in cohort reports the total numbers in the cohorts born 2008, 2009 and 2010 from which the sample for modeling was drawn.

(4) 3% for A; 8% for B; 5% for C; 1% for D and E. The threshold for each is selected to approximate the actual percentage with the outcome targeted by age 2.

9. Sensitivity to agency data drawn on

134 One of the tasks of the feasibility study is to assess which agency data would be needed in order to make the best early identification predictions. Table 11 examines the impact on predictive accuracy and selected measures of proportionality of either adding or successively dropping agency data:

- column B presents the base model developed for the 2010 cohort and simulates scoring the 2007 cohort and referring the most risky 3,000 children identified by the model

- column A presents a model that is applied to the same population as the base model[23] but also draws on linked Ministry of Health data

- column C drops from the base model the linkage with births data and presents a model that scores only children identified through their inclusion in a caregiver's benefit, and relies on predictors sourced from benefit, care and protection and Corrections administrative data - this removes around three quarters of children (being those not included in benefit within three months of birth) from the population screened

- column D also drops the sentencing administrative data, and instead includes a predictor variable that shows whether transitions between benefit and prison occurred in the previous five years (derived from benefit data) – this final model scores only children identified through their inclusion in a caregiver's benefit, and relies only on predictors sourced from benefit and care and protection administrative data.

Sensitivity to a more extensive data linkage - the inclusion of Ministry of Health data

135 Additional variables considered in the model that includes Ministry of Health data are outlined in Box 2 and described in more detail in the technical companion report. Inclusion of Ministry of Health data resulted in large statistically significant improvements in the predictive value of the mental health and address change markers and a small improvement in that of the low birth weight or pre-term marker, as measured by the AUR for a model that included only that variable (but no significant improvement in the predictive value of the parenting demands and family violence markers).

136 Despite this, the model that includes Ministry of Health data performs no better than the base model comparing AUR and the misclassification rate at the three percent threshold. Performance as measured by cumulative lift is slightly improved (from 8.3 to 8.7 percent), indicating that at highest risk scores the model with Ministry of Health data slightly improves the identification of children with findings of maltreatment. When used to risk score the 2007 study cohort and to identify the 3,000 children with the highest risk scores, however, sensitivity and specificity are unchanged compared with the base model, and positive predictive value shows only a very small improvement.

137 On selected measures where the base model shows disproportionality relative to known maltreatment (see Tables 7 and 8), the model that includes Ministry of Health data would reduce but not fully address disproportionality (with lower representation of Māori children, children of single parents, and children supported by benefits). The model would leave largely unchanged over-representation of new-borns where other children have a care

---

[23] Numbers and the rate of findings differ slightly from the base model because with the addition of Ministry of Health identity data, some linkages were re-formed.

and protection history, and would intensify the over-representation of those with a parent or caregiver with a care and protection history as a child (Table 11).

Box 2 Additional variables considered in the model incorporating Ministry of Health data

MoH - low birth weight or pre-term
(yes - <2.5Kg or <37 weeks gestation / no or unknown)
Based on
- Birth registration data
- MoH Maternity Collection data

MoH - parenting demands
(high demands - multiple birth child, child with a birth abnormality, other children under 2, or 3 plus other children / no other children / other children but not high demands)
Based on
- Birth registration data
- MSD benefit data on other children included in benefit
- MoH hospitalisation data on birth abnormalities and number of previous hospital births

MoH - Family violence
(no events / notification or contact record in 1 of the last 12 months / notification or contact record in 2+ of the last 12 months, or mother hospitalised for assault in the last 5 years, or other child hospitalised for maltreatment, intracranial or long-bone fracture injury by age 1)
Based on:
- MSD CYF data on Police family violence notifications and contact records for other children identified through benefit and birth records
- MoH hospital admission data for mother
- MoH hospital admission data for other children identified through benefit and birth records

MoH – 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)[24]
Based on:
- MSD benefit data on incapacity codes recorded when parents or caregivers have claimed Sickness or Invalid's Benefits
- MoH data on maternal history of addiction and non-addiction mental health service usage
- MoH data on maternal history of prescribed pharmaceuticals used solely to treat substance use and other mental health disorders
- MoH data on maternal history of clinical codes indicating substance abuse and other mental health disorders recorded for hospital admissions

MoH – parents' or caregivers' number of address changes
(no address changes / 1-2 address changes / 3 plus address changes / missing)
Based on:
- MSD benefit data on addresses recorded in benefit data in the last year
- MoH data on meshblocks recoded in Primary Health Organisation data for mother in the last year
- MoH data on domicile codes recoded in hospitalisation data for mother in the last year

MoH - mother a smoker
(yes / no or unknown)
Based on MoH hospitalisation data on maternal smoking status at child's birth

MoH - other child not fully immunised and no non-consent at age 1
(yes / no or unknown)
Based on MoH National Immunisation Register data for other children identified through benefit and birth records. Coded "yes" where another child has immunisation status (excluding pneumococcal conjugate vaccine) "not complete" at ages 6, 8 and 12 months and there is no record of a parent or guardian declining immunisation.

---

[24] Substance abuse is prioritised so that this category applies where a person is recorded as having both a substance abuse and another mental health disorder, or where one parent has a substance abuse disorder and another has a mental health disorder.

Table 11. Models based on more extensive and more restricted linkages of data, 2010 cohort check base model

Stepwise logistic regression models developed for children born in 2010 with a birth registration or seen on benefit by 3 months of age (A and B) or children seen on benefit by 3 months of age (C and D) and used to risk score children born in 2007[1]

	A	B	C	D
	Ministry of Health, benefit, care and protection, sentencing and birth data)	Base model (benefit, care and protection, sentencing and birth data)	benefit, care and protection and sentencing data	benefit and care and protection data
Number in cohort	63,200	63,176	16,931	16,944
% with findings of maltreatment by age 2	2.7	2.8	8.4	8.4
Sample size for modelling[2]	8600	8750	16,931	16,944
AUR (validation sample)	88%	89%	74%	74%
(95% confidence interval)	(87%-90%)	(88%-91%)	(71%-76%)	(72%-77%)
Misclassification rate at selected threshold[3] (scored 2007 study population)	4.1%	4.2%	12.4%	12.2%
Cumulative Lift at selected threshold[3] (scored 2007 study population)	8.7	8.3	2.6	2.8
Using the 2010 model to risk score the 2007 cohort and the 3,000 children with the top scores to define high-priority for referral:				
% referred children with findings by age 5 (PPV)	30.8	30.7	31.3	31.0
% all children with findings by age 5 referred (sensitivity)[4]	30.7	30.7	30.9	31.2
% all children with no findings by age 5 not referred (specificity)[4]	96.5	96.5	96.5	96.5
Profile of referred children (%):				
Māori child (of non-missing)	66.1	69.2	67.7	66.8
Other children with a care and protection history in the last 5 years	58.2	59.9	56.2	57.3
Single parent	78.6	87.7	89.5	86.0
Mother or caregiver aged under 25	57.3	54.6	61.2	61.0
Parents or caregivers with a care and protection history as a child	69.8	57.2	54.2	53.7
Child seen on benefit by 3 months of age	85.7	92.3	100.0	100.0

Notes:

(1) Based on conservatively linked data. This is known to understate the proportions with CYF contact and findings of maltreatment.

(2) In Columns C and D the sample size for modeling is the same as the number in the cohort study population. This is because with a higher proportion with findings of maltreatment by age two, it is not necessary to apply sampling to over-sample children who had the outcome targeted.

(3) 3% for A and B; 8% for C and D. The threshold for each is selected to approximate the actual percentage with findings of maltreatment by age 2.

(4) To allow comparison across models, sensitivity and specificity are calculated using the total number of children in the study population with and without substantiated findings by age 5 calculated for the base model study population as the denominators for all models. (They are not calculated only for the population screened in the particular model, which is lower in models C and D).

Sensitivity to more restricted data linkages

138 Models based on the more restricted data linkages (D and C) have a lower AUR, lower cumulative lift and a higher misclassification rate than the base model. This is largely a mechanical effect – the base model targets an outcome that is more rare in a relatively low-risk population, and this results in fewer false positive predictions and higher cumulative lift.

139 Comparing positive predictive value and sensitivity when identifying the 3,000 children with the highest scores, the more restricted data linkages make only a small difference to sensitivity and positive predictive value, and models C and D differ little from one another in their predictive accuracy.

140 A clear limitation of these models is that they are only able to identify risk for new-born children known to MSD.

141 Compared to a model based on benefit and care and protection data only:

- adding data on Corrections managed sentences provides no substantive gain

- adding births data also provides no substantive gain in predictive accuracy as measured by sensitivity, specificity and positive predictive value, but does go some way to addressing the disproportionate representation of children supported by benefits among those identified as at high risk, and provides the basis for the deployment of separate models applied across the population to fully control disproportionality relative to known maltreatment on this or other dimensions of concern, and for monitoring and controlling disproportionate representation of Māori children

- beyond the incorporation of births data, the further addition of Ministry of Health data provides very little gain in predictive accuracy, and only modest benefit in reducing disproportionality on selected dimensions.

10. Main predictors

142 The administrative databases drawn upon for the study do not provide full coverage and capture of all known risk and protective factors for maltreatment. They do not, for example, capture risk factors such as poor bonding between mother and child or unrealistic parental expectations. Nor do they capture important protective factors such as the presence of a supportive adult. This prevents any investigation of the data aimed at identifying the most important risk or protective factors, or explaining the causes of maltreatment and sources of resilience that protect children who appear at high risk from harm.

143 However, stepwise logistic regressions do allow the predictive utility of different administratively derived risk markers to be assessed. This modelling approach can produce unstable variable selection (a problem that is not a concern where the task is prediction - as in the case of this study - rather than hypothesis testing (Austin and Tu, 2004)). However, we find that the most useful predictors are relatively stable across birth cohorts and across various sensitivity tests.

144 Figure 3 looks across 16 of the models developed for the research (four base models for different birth cohorts, six separate models for different population sub-groups, four models for alternative target variables, and the two models that consider less extensive linkages of data). It plots the frequency with which different predictor variables appear in the models, and the average step-wise ranking in models in which they appear (a rank of one indicating that the variable has the highest predictive value in the model).

145 The most frequently appearing and most important predictors are:

- the presence of other children with contact with care and protection services in the previous five years[25]

- the length of time the mother or primary benefit recipient with care of the subject child was supported by benefit in the previous five years

- having a parent or caregiver who was known to have had contact with care and protection services in their own childhood.

146 These three variables are also the top three predictors in the model that includes Ministry of Health data (other variables appearing in that model are the smoking status of the mother at birth, and the mental health history and address change indicators that include Ministry of Health data).

147 The importance of the remaining predictors is more variable across the 16 models assessed in Figure 3, with indicators related to mental health, sentencing history, family violence, single parent status and caregiver age appearing the next most frequently in the models, and the mental health indicator, CYF service centre, and family violence, single parent status, caregiver age and sentencing history indicators having the next lowest average ranking (highest predictive value) in the models in which they appear.
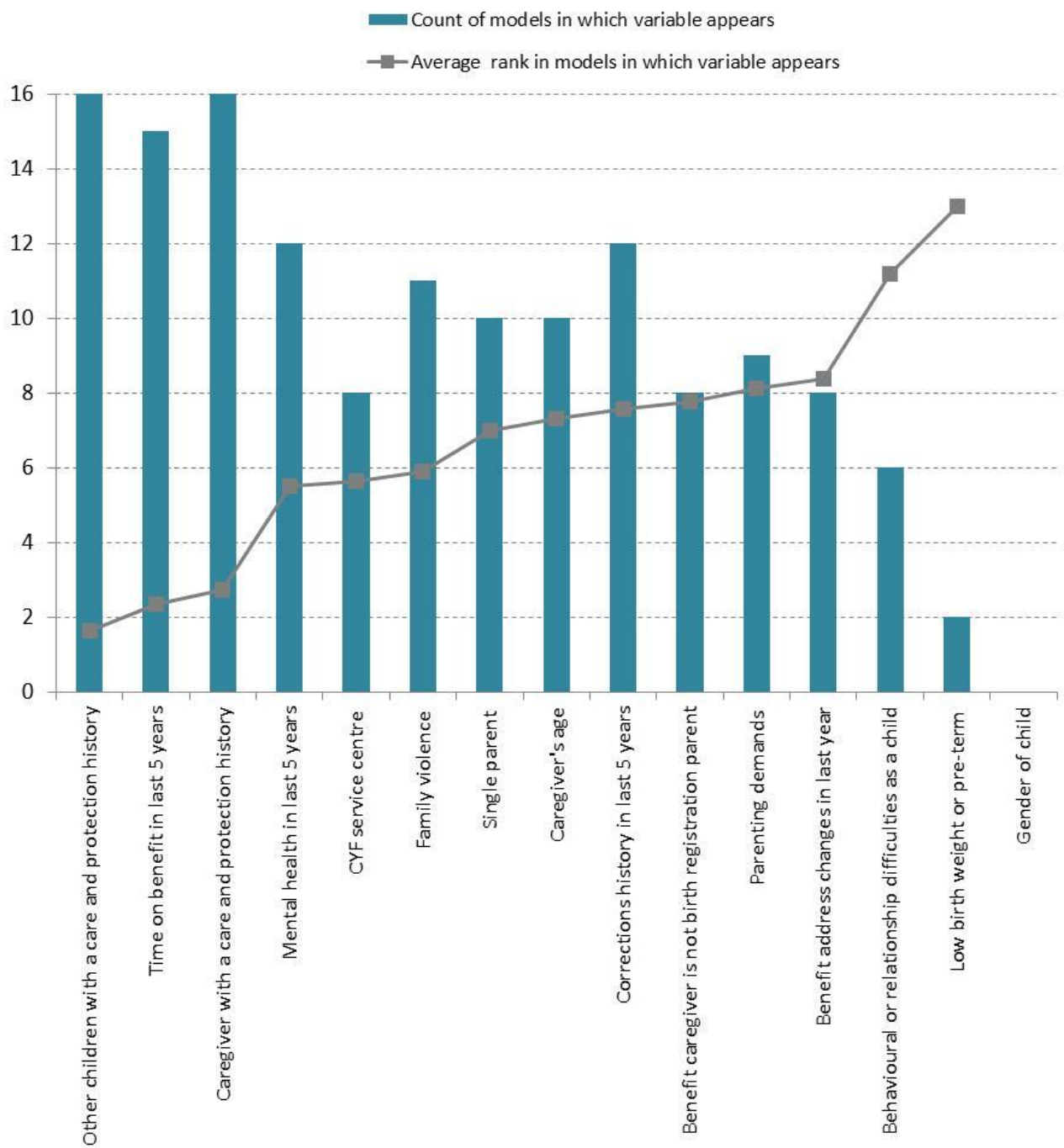
148 As noted, these variables cannot be interpreted as factors that cause maltreatment to occur, only as useful predictors for identifying some of the children at high risk, whose families may benefit from preventive services (Wu et al., 2004).

---

[25] Excluding cases where other children had Police family violence notifications but no other contact with care and protection services.

149 Of all the variables considered, data on whether the child was low birth weight or pre-term and the gender of the child had the least predictive utility.
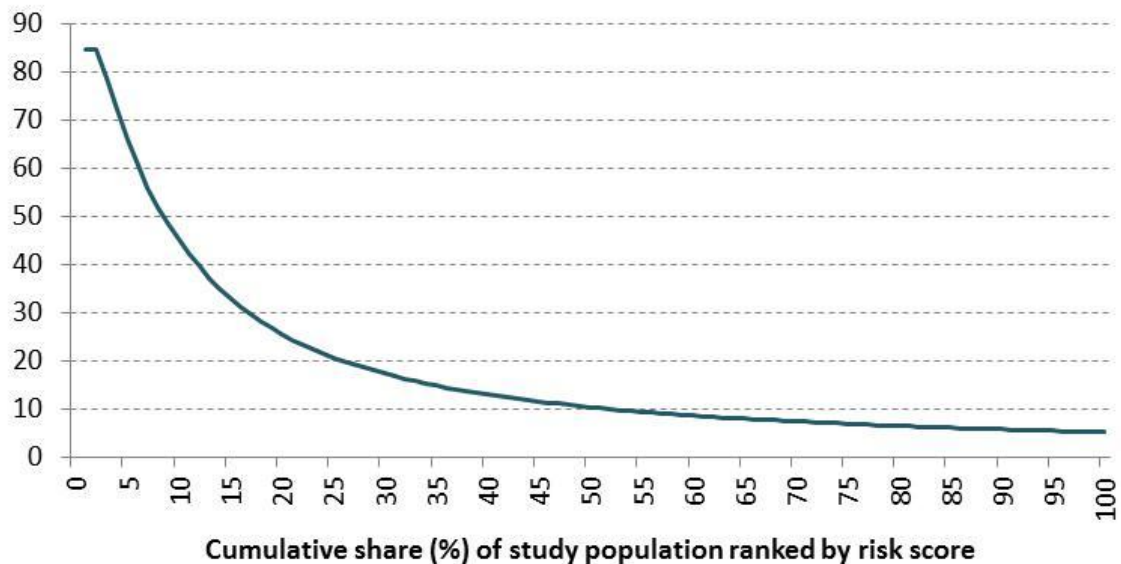
Figure 3. Main predictor variables across 16 models[1]

Notes:
(1) Obtained from stepwise logistic regression models where the significance entry level was set to 1 and significance stay level was set to p<.02. Where a variable was not included in a model but was the criterion on which separate models were developed, it is counted as appearing in the model.

150 Figure 4 takes the 11 variables with the highest predictive value (excluding CYF service centre) and uses these to create a series of 11 risk markers. Among children with the highest two percent of risk scores, 84.7 percent have at least five of the 11 risk markers. Among children with the highest five percent of risk scores, the proportion is 66.0 percent. Among children in the study population overall, the proportion is 5.2 percent.

Figure 4. Percent of 2007 study population with five or more out of 11 selected risk markers[1] by risk score

2010 base model applied to children born in 2007 with birth registration or on benefit by 3 months of age

Cumulative share (%) of study population ranked by risk score

Notes:

(1) Selected risk markers are:
- the presence of other children with contact with care and protection services in the previous five years
- mother or primary benefit recipient with care of the subject child was supported by benefit for more than 80 percent of the previous five years
- a parent or caregiver who was known to have had contact with care and protection services in their own childhood
- a parent or caregiver with a benefit for substance abuse or other mental health disorder in the last five years
- the presence of other children known to CYF for whom there are Police family violence notifications or contact records in the last year
- a parent or caregiver who is single
- mother or primary benefit recipient aged under 20
- a parent or caregiver with a Corrections sentence in the previous five years
- a benefit caregiver who is not a parent listed on the birth registration
- a parent or caregiver with high parenting demands
- a parent or caregiver with at least one change in benefit address in the last year.
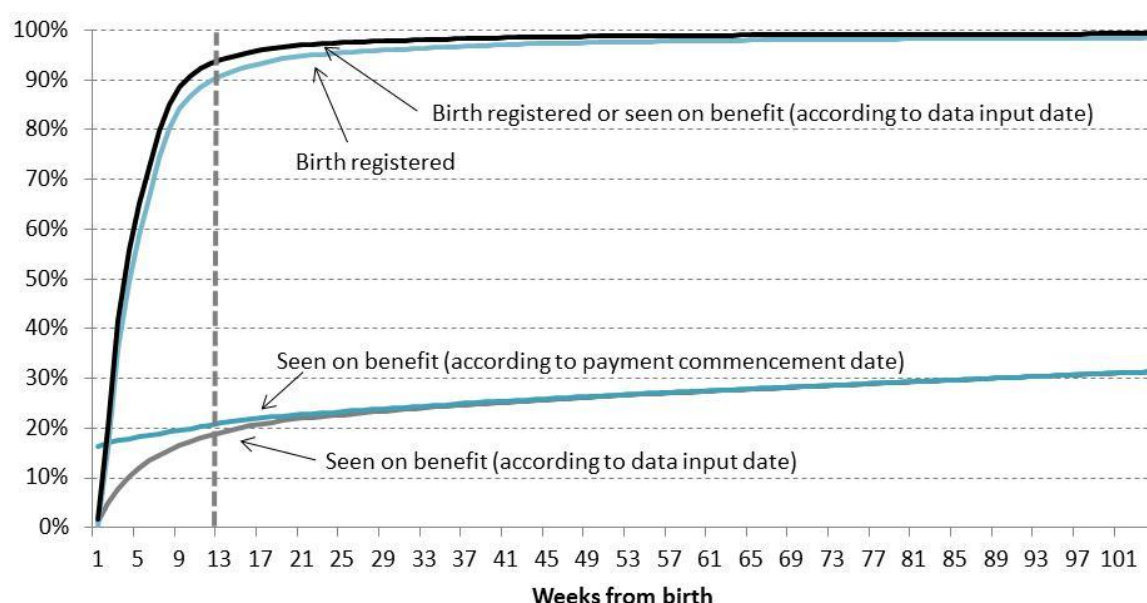
11. Time to first observation in birth registration and benefit data

151 One of the implications of drawing on administrative data is that time lags in the generation of those data would constrain the timeliness of early identification via a PRM tool. This constraint does not apply in an early identification model that relies on professional judgement, where children are known to professionals.

152 Figure 5 shows the proportion of New Zealand-born children who would be observed in an administrative data base with increasing time from birth, based on less conservatively linked data for children born in 2007.[26] (The less conservative data linkage is used because it provides the best estimate of the total number of unique children and the degree to which they appear in multiple administrative systems – with a conservative data linkage a child might appear in separate systems but not be recognised as the same child).

153 In many cases, the claim for the child's inclusion in benefit occurs some weeks after the birth, and the payment commencement date is then backdated to the date of birth. The research data available for this study provides only the payment commencement date. Figure 5 links in other data to provide an estimate of the earliest date the child would actually become known to the system (according to the data input date for the claim for inclusion).

Figure 5. Percentage of New Zealand-born children observed in birth registration and benefit data with increasing weeks of birth
2007 birth cohort [1]

Notes:
(1) Based on less conservatively linked data which provides the best estimate of the number of unique children. The total number of New Zealand-born children in the cohort is estimated to be all children with either a New Zealand birth registration by June 2012, seen on benefit by 3 months of age or with a Ministry of Health maternity record.

154 Based on the trajectories of the 2007 cohort, by three months of age (13 weeks):

- just under one in five children would be known to the benefit system
- 91 percent would have their birth registered

---

[26] Time lags for Ministry of Health data are not considered as these data provide little gain in predictive accuracy.

- 94 percent would be known to the benefit system or have their birth registered.

155 Children known to the benefit system or with their birth registered by three months of age accounted for 87 percent of all the New Zealand-born children in the 2007 cohort with substantiated findings of maltreatment by age five.[27]

156 Extending the window to include children seen on benefit or with a birth registered by six months of age would boost coverage of children overall to an estimated 98 percent, and increase coverage of children with substantiated findings of maltreatment by age five to 94 percent.

157 If there is a decision to proceed with the trialling of a PRM tool, additional work could be undertaken to investigate the feasibility of:

- applying a model in the ante-natal period (using, for example, administrative data on receipt of welfare benefits for reason of pregnancy)

- using birth notification data as the basis for screening – a hospital (or doctor or midwife present at a birth) is required to notify the Registrar General of Births Deaths and Marriages of a birth within five days of the child being born – use of this notification data would provide very high coverage and increase timeliness.

158 Potential for unintended consequences (for example, dissuading uptake of maternity services by high risk mothers) would need to be considered.

---

[27] Based on less conservatively linked data which provides the best estimate of the number of unique children and their contact with CYF. The total number of New Zealand-born children in the cohort is estimated to be all children with either a New Zealand birth registration by June 2012, seen on benefit by 3 months of age or with a Ministry of Health maternity record.

12. Legal and operational feasibility

Legal feasibility

159 Data linkage for the feasibility study reported here falls under research exceptions to the New Zealand Privacy Act given that the information is being used for research purposes and will not be published in any form that could reasonably be expected to identify the individuals concerned.

160 Information sharing and data linkage for fully implemented PRM requires separate consideration.

161 Linkage of benefit and care and protection data would comply with the Privacy Act as it involves use of data collected for one purpose for a directly related purpose, and does not require separate authorisation. However formalisation of this within-agency data linkage for the purpose of PRM could be a useful mechanism for avoidance of doubt.

162 Linking data held by different Government agencies would require separate authorisation (such as an Approved Information Sharing Agreement) where it does not involve use of data collected for one purpose for a directly related purpose.

Operational feasibility

163 MSD has already developed and operationalised PRM tools that identify benefit recipients at high risk of long-term benefit receipt. This previous work has established that operationalising PRM tools is technically feasible. It also provides some lessons that could usefully inform any operationalisation of a new model. These include the importance of early engagement with front-line delivery agents, training, and feedback mechanisms, and the need to link in death registration data.

164 Operationalising a PRM would require systems that ensure the secure handling of the data drawn on and generated, and for ensuring protection from mis-use.

165 A Privacy Impact Assessment would be a useful first step in any implementation. It would provide a framework for privacy protection in systems that feed in data from contributing agencies and databases, that link data, that apply and regularly review risk scoring algorithms and that generate and disseminate PRM information.

166 PRM algorithms would need to be refreshed at regular intervals. They could, for example, be updated annually to reflect the most recent data with a more detailed review every two or three years.

13. Limitations

167 Against the strengths of the administrative data drawn upon (Putnam-Hornstein et al., 2013a; Brownell and Jutte, 2013), and the strengths of this approach to prediction, there are a range of limitations. Most of these have been discussed, and many apply not only to PRM but also to other approaches to risk assessment.

- Combining data from across social sector administrative systems requires linking on name, date of birth and other potentially identifying information. Not all records relating to the same individuals are able to be linked. While we seek to avoid false positive matches by taking a conservative approach, some may be made, and an inevitable consequence of this approach is that some matches are missed.

- An implication of missed matches is that models will be "trained" with, in some cases, imperfect information on whether a child went on to have findings of maltreatment. Predictions will miss some high priority children as a result.

- Reporting and recording errors are likely to be present in the administrative data, and errors may also be made in assembling variables for analysis. Although efforts are made to check and correct for errors, not all are able to be identified and accounted for.

- Care and protection data are known to understate the lived experience of maltreatment (based on comparison with studies of self-reported maltreatment) (Gilbert et al., 2009; Runyan et al., 2002). Predictions formed based on an incomplete picture of the outcome of concern might be expected to understate risk.

- Analytical use of unit-level care and protection administrative data is emergent, and some aspects of the quality of these data are still being understood and enhanced.

- Māori children, children from poor communities, and children whose families are supported by benefit and/or are already known to care and protection services are over-represented among those who have maltreatment substantiated (Tables 4 and 5). To a large extent, this will reflect a higher incidence of risk factors for maltreatment in these populations, but there is also a possibility that it partly reflects bias as a result of, for example, higher rates of surveillance for maltreatment in poor communities or higher rates of classifying a given event as maltreatment where a family has previous contact with care and protection services (Leventhal, 1988). With this in mind, the impact of added disproportionality in the predictions made by PRM in any implementation needs to be acknowledged and mitigated to the extent possible.

  The research data assembled for this study provide new opportunities for addressing the question of whether bias exists (using the approach applied in Drake et al., 2011, for example). This is the focus of further research currently under way.

- Measured contact with the different administrative systems may have different meanings at different times depending on policy settings and administrative practice, and great care is needed in interpretation of changes over time (Gilbert et al., 2011) and in ensuring that any PRM driven off administrative data is responsive to change. Emotional abuse findings, in particular, have been sensitive to procedural changes.

- The administrative nature of the data means that it does not provide full capture of the risk and protective factors for maltreatment, and in some cases the degree to which the available information can accurately measure or proxy risk or protective factors is limited. The main predictors highlighted in this report cannot be interpreted as factors that cause maltreatment to occur.

- In addition to failing to identify children as high priority as a result of errors in data linkage and false negative predictions at the threshold applied, a PRM tool applied at birth would miss children not born in New Zealand who move here during their early years and are maltreated.

- At any threshold, some children who would not go on to have the outcome of concern, (including children who appear to be high risk based on the historical variables used by the PRM tool but whose families have resolved their problems), would be referred.

- A tool based on the circumstances of a child at birth misses newly emerging circumstances that threaten or protect the child as he or she ages (Leventhal, 1988). If PRM is to be implemented, the feasibility of additional screening based on information that becomes available at older ages could be explored.

- If interventions become tied to predictions of risk, associations between risk markers and outcomes will alter (Leventhal, 1988), and may need to be adjusted for over time.

14. Discussion

168 PRM tools show promise in helping to identify early some of the children at highest risk of maltreatment. They offer the promise that, with sensitive out-reach, more high-needs families can be engaged in effective early intervention services that they experience as helpful and, as a consequence, children's exposure to maltreatment and its associated long-term adverse impacts on health and wellbeing can be reduced.

169 The efficacy and cost effectiveness of PRM as a screening tool is dependent on it being coupled with early intervention service responses that reduce maltreatment.

170 There is evidence that high quality home visitation programmes for high-risk families are successful in reducing harsh parenting and reducing the proportion of children who are notified to care and protection services or hospitalised as a result of injury, but not all home visitation programmes have this effect (Nelson et al., 2013; Moyer, 2013; MacMillan et al. 2009; Fergusson et al., 2012). Provision of information resources and social work services via primary health care shows promise (Dubowitz et al. 2009; Dubowitz et al. 2012). High quality early childhood education programmes that emphasise parent involvement warrant consideration (Mersky et al., 2011). Interventions designed to improve the community environment in which children are raised also show promise (Daro and Dodge, 2009).

171 In the New Zealand context, the success of early intervention services depends on the extent to which they are successful in engaging and improving outcomes for Māori children and their whānau (Cooper and Wharewera-Mika, 2009; Cram, 2012; Durie et al., 2010; Fergusson et al., 2012; Tamati et al., 2008).

172 The Whānau Ora approach to social service delivery introduced in 2010 offers a culturally anchored approach that "empowers whānau as a whole – rather than focusing separately on individual family members and their problems – and requires multiple government agencies to work together with families rather than separately with individual relatives" (Durie, 2013). This builds on earlier developments in the delivery of health services, and offers opportunities for new ways of working with whānau to improve the lives of their children (Cram, 2012; Durie et al., 2010; Boulton et al., 2013).

173 At the same time, the evaluation of the Early Start home visitation programme and the Incredible Years pilot study suggests that (with attention to culturally appropriate delivery in the case of Early Start), "mainstream" services can also improve quantitatively measured outcomes for Māori and are experienced positively by Māori parents and caregivers (Fergusson et al., 2012; Sturrock and Gray, 2013).

174 Efficacy (and ethical justification) is also dependent on there being capacity to meet the demand for early intervention services generated by the application of PRM tools. It is unethical to provide screening where there is no accepted intervention that will have a positive impact on the condition screened for, or where there is a suitable intervention but there is insufficient capacity to address the needs identified (Wilson and Jungner, 1968). The ethical review discusses these and other considerations.

175 If the decision is to proceed, trialling will be required to establish whether use of PRM together with an early intervention response reduces child maltreatment, and the cost effectiveness of the approach. Assessment of the cost effectiveness of PRM will need to take account of the cost of screening, and the cost of early intervention services.

176 A question that will not be able to be answered unless the approach is trialled is whether PRM identifies new-born children who would not have in any case been referred and prioritised in early identification and intervention efforts.

177 There is some suggestion from results to date that some children identified through PRM might not be otherwise identified as early.

- When defining as high priority the 3,000 or five percent of children with the highest risk scores, three in ten high-priority children had a caregiver with changes in benefit address in the previous year based on 2007 data (Table 5) – raising the possibility that if they were known in one location, they might not be in their new location.

- While most children who would have been scored as high priority at this threshold were in families with children previously known to care and protection services, for the 2010 and 2011 cohorts it is estimated that in at least half of cases, the new-born child themselves was not already known to Child, Youth and Family and had not already been referred to the Family Start home visitation programme by three months of age.[28]

---

[28] Based on Child Youth and Family notification data, data on whether there was an open CYF phase of engagement for the child or for other children in the family when the child was born, and linked Family Start referral data. Estimates are adjusted to take account of undercounts of agency contact as a result of inability to link data (in most cases due to missing or approximate identity details). Estimates are for children in sites where Family Start is available. The proportion with no prior agency contact will be higher in sites where Family Start is not available.

15. Conclusions

178 Based on the findings of this study, PRM tools based on linked administrative data could be used to identify early some of the new-born children at high risk of maltreatment. Compared to other models reviewed in the international literature, these models perform well, both overall, and for Māori children.

179 Not all children who go on to experience substantiated maltreatment would be able to be identified early using this approach (depending on the threshold for referral chosen, most may not be able to be identified). This is because:

- in some cases information that should inform the assessment of risk is not available from administrative data, or is not able to be taken into account because of difficulties with data linkage

- as with any risk prediction tool, at any threshold for referral, PRM tools will inevitably identify as low risk some children who go on to experience abuse or neglect, and identify as high risk some children who do not

- it is inherent in the approach that children who experience maltreatment but about whom less is known because the family has little or no prior contact with social agencies will be under-represented among those identified as high risk.

180 It is recommended that, if taken to trial, PRM tools should therefore not be the sole mechanism for identification of children at high risk, and should complement rather than replace professional judgement.

181 If the decision is to proceed to trial, it is recommended that:

- benefit and care and protection data be the basis of the modelling and that inclusion of birth registration or notification data and death registration data[29] also be considered to allow:

    - risk to be considered across the whole of the population

    - separate models to be deployed to control disproportionality on dimensions of concern

    - monitoring and control of over-prediction relative to known maltreatment for Māori children

- given the potential for error in data linkage, a system be put in place for review of the linkages that inform the risk scoring before details of a high risk child are released as part of any working PRM - having an effective system of review in place may allow a less conservative approach to data linkage to be taken, and improvements in predictive accuracy to be realised.

182 Careful thought would need to be given to potential uses of PRM risk scores and the underlying information, and the role they play in risk assessment and decision making. In addition to directly generating early identification referrals, high PRM risk scores could be made available to front-line professionals to inform their decisions about whether to refer a child about whom they have concerns to care and protection or other services.

183 These potential uses would require careful consideration. Reliance on a tool can mean giving insufficient weight to risks and protective factors that can be directly observed, but

---

[29] Birth weight and gestation, which are items of Ministry of Health information passed to the Registrar General of Births, Deaths and Marriages as part of the birth registration, would not be required as these contribute only marginally to predictive accuracy.

are not able to be captured by the tool. It would be important that any front-line professionals who were to have access to risk score information received training in how to interpret and apply that information, and on the circumstances in which it should be over-ridden (Gambrill and Schlonsky, 2000).

184 Because the administrative data that are able to be collated and linked provide partial and selective capture of risk information, a cautious approach would need to be taken to any dissemination of the information that sits behind the risk score for an individual child.

185 One potential risk is that decision making could be made more rather than less resource intensive and complex if having access to the administrative information that informs the scores means front-line professionals need to spend time reconciling this information with what they are able to observe. Information on parental and sibling care and protection history is a case in point.

- In some cases, it will include history that front-line professionals are unable to observe (because, for example, PRM has linked records held under an alias) or to confirm (because, for example, parents are not aware of past notifications and findings recorded for them or their children).

- In other cases, front-line professionals will be aware of history that PRM information will not include (because, for example, parental history occurred prior to the collection of electronic records, or is not able to be readily linked because of imprecision in identity recording).

16. Next Steps

186 Findings from this feasibility study, in combination with those from the ethical review, will inform decisions around next steps in relation to the possible trialling of PRM.

187 Beyond this study, the data assembled provide the basis for a wide range of new analysis that can aid the understanding of, and response to, child maltreatment (Putnam-Hornstein et al., 2013a; Brownell and Jutte, 2013).

188 Predictions from the PRM models developed in this study could be compared with those from a model that quantifies and scores according to the accumulation of risk factors (Larsen et al., 2008; Linder and Sexton, 2011). New ways of describing the data (for example using cluster analysis techniques to profile sub-groups of children who are known to experience maltreatment or are identified as high risk), could also be explored.

189 Research is under way to better understand Māori children's disproportionate representation in maltreatment statistics, and to explore more generally whether bias exists in measured contact with care and protection services.

190 Finally, with further data linkage to incorporate information on the availability and use of services (teen parent education units, parenting programmes and home visitation services for example), the data offer new opportunities for analysis to build understanding of the impact of these services on outcomes for children and their families and whānau.

References

Austin, P.C. and Tu, J.V. (2004). Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality *Journal of Clinical Epidemiology,* 57, 1138–1146

Barlow, J., Fisher, J.D. and Jones, D. (2012). *Systematic Review of Models of Analysing Significant Harm*. (Research Report DFE-RR199). London: DfE [online]. Available from: https://www.education.gov.uk/publications/standard/publicationDetail/Page1/DFE-RR199

Bartholet, E., Wulczyn, F., Barth, R.P., & Lederman, C. (2011). *Race and Child Welfare.* Chicago: Chapin Hall at the University of Chicago.

Baumann, D.J., Dalgleish, L., Fluke, J., Kern, H. D. (2011). *The Decision Making Ecolog*y. American Humane Association: Washington, DC.

Boulton, A., Tamehana, J. and Brannelly, T. (2013). Whānau-centred health and social service delivery in New Zealand: the challenges to, and opportunities for, innovation. *MAI Journal*, 2(1), 18-32.

Browne, K. and Chou, S. (undated). *A Literature Review on Systems for Early Prediction and Risk Detection in Child Protection in Europe.* Deutsches Jugendinstitut. Available from: http://www.dji.de/bibs/Expertise_Browne.pdf

Brownell, M.D. and Jutte, D.P. (2013). Administrative data linkage as a tool for child maltreatment research *Child Abuse & Neglect* 37, 120–124.

Cameron, M.P., Cochrane, W., Gordon, C., and Livingston, M. (2013). *The locally-specific impacts of alcohol outlet density in the North Island of New Zealand, 2006-2011.* Research report commissioned by the Health Promotion Agency. Wellington: Health Promotion Agency

Centre for Social Research and Evaluation. (2012). *Children's Contact with MSD Services. Ministry of Social Developmen*t. Available from: http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/childrens-contact-with-msd-services/index.html

Chabot, M., Fallon, B., Tonmyr, L., Maclaurin, B., Fluke, J., and Blackstock, C. (2013). Exploring alternate specifications to explain agency-level effects in placement decisions regarding aboriginal children: Further analysis of the Canadian Incidence Study of Reported Child Abuse and Neglect Part B. *Child Abuse & Neglect* 37, 61-76.

Child Youth and Family (2006). *Children at increased risk of death from maltreatment and strategies for prevention*. Ministry of Social Development.

Choi, J. M., (2010). *A selective sampling method for imbalanced data learning on support vector machines.* PhD Dissertation. Iowa State University, USA. Available online *lib.dr.iastate.edu/cgi/viewcontent.cgi?article=2549&context=etd*

Cooper, E., and Wharewera-Mika, J. (2009). *Māori Child Maltreatment: A Literature Review Report*. Auckland, New Zealand, Te Kahui Mana Ririki. Available from: http://www.ririki.org.nz/index.php/research

Cram, F. (2012). *Safety of subsequent children - Māori children and whānau – A review of selected literature*. Families Commission. Available from: http://www.familiescommission.org.nz/web/māori-children-whānau/index.html

Curtis, M. (2012). *Statistical Analysis and Summary of Themes Family Violence Death Reviews of Deaths between 2004 – 2011*. New Zealand Police.

http://media.nzherald.co.nz/webcontent/document/pdf/201239/family-violence-death-review-2004-2011.pdf

Daro, D. and Dodge, K. (2009). Creating Community Responsibility for Child Protection: Possibilities and Challenges. *The Future of Children*, Volume 19, Issue 2, 67-93.

D'Andrade, A., Austin, M., and Benton, A. (2008). Risk and Safety Assessment in Child Welfare: Instrument Comparisons. *Journal of Evidence-based Social Work*, Volume 5, Issue 1-2, 31-56.

Drake, B. (1996). Unraveling unsubstantiated. *Child Maltreatment*, Volume 1, Issue 5, 261-271.

Drake, B., Jolley, J., Lanier, P., Fluke, J., Barth, R., and Jonson-Reid, M. (2011). Racial Bias in Child Protection? A Comparison of Competing Explanations Using National Data. *Pediatrics* Volume 127, Issue 3, 471-478.

Dubowitz, H., Feigelman, S., Lane, W., and Kim, J. (2009). Pediatric Primary Care to Help Prevent Child Maltreatment: The Safe Environment for Every Kid (SEEK) Model. *Pediatrics* Volume 123, 858-864.

Dubowitz, H., Lane, W.G., Semiatin, J.N., and Magder, L.S. (2012). The SEEK model of pediatric primary care: can child maltreatment be prevented in a low-risk population? *Academic Pediatric Association,* Jul-Aug Volume 12, Issue 4, 259-68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22658954

Durie, M., Cooper, R., Snively, S., Grennell, D., and Tuaine, N. (2010). Whānau Ora: Report of the Taskforce on Whānau-Centred Initiatives. Wellington: Ministry of Social Development. Available from: http://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/planning-strategy/whānau-ora/whānau-ora-taskforce-report.pdf

Durie, M. (2013). *Whānau Ora Fact Sheet: Paenga-whāwhā / April 2013*. Wellington: Te Puni Kōkiri Available from: http://www.tpk.govt.nz/en/in-focus/whānau-ora/

Fallon, B., Chabot, M., Fluke, J., Blackstock, C., Maclaurin, B., and Tonmyr, L. (2013). Placement decisions and disparities among Aboriginal children: Further analysis of the Canadian incidence study of reported child abuse and neglect part A: Comparisons of the 1998 and 2003 surveys. *Child Abuse & Neglect,* 37, 47-60.

Fergusson, D., Boden, J. and Horwood, J. (2012). *Early Start Evaluation Report: Nine Year Follow-up*. Wellington: Ministry of Social Development. Available from http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/evaluation/early-start/index.html

Fluke, J., Shusterman, G., Hollinshead, D., and Yuan, Y.T. (2008). Longitudinal analysis of repeated child abuse reporting and victimization: multistate analysis of associated factors. *Child Maltreatment*, 13 (1), 76 – 88.

Fluke, J. (2009). Allegory of the cave: On the theme of substantiation. *Child Maltreatment*, 14, 69-72.

Gambrill, E., and Shlonsky, A. (2000). Risk Assessment in Context. *Children and Youth Services Review*, Volume 22, Issues 11–12, 813-837*.*

Gilbert, R., Widom, C.S., Browne, K., and Fergusson, D., and Webb, E. (2009). Burden and consequences of child maltreatment in high-income countries. *The Lancet* 373.9657 (Jan 3-Jan 9): 68-81.

Gilbert, R., Fluke, J., O'Donnell, M., Gonzalez-Izquierdo, A., Brownell, M., Gulliver, P., Janson, S. and Sidebotham, P. (2011). Child maltreatment: variation in trends and policies in six developed countries. *The Lancet* December (online).

Harcourt, B. E, (2006). *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. University of Chicago.

Hussey, G., Marshall, J., English, D., Knight, E., Lau, A., Dubowitz, H., and Kotch, J. (2005). Defining maltreatment according to substantiation: Distinction without a difference? *Child Abuse and Neglect,* Volume 29, Issue 5, 479-492.

Kerslake Hendricks, A. and Stevens, K. (2012). *Safety of Subsequent Children International literature review.* Wellington: Families Commission. Available from: http://www.familiescommission.org.nz/publications/research-reports/safety-of-subsequent-children

Larson, K., Russ, S.A., Crall, J.J., and Halfern, N. (2008). Influence of Multiple Social Risks on Children's Health. *Pediatrics* 121(2), 337 -344.

Linder, S.H. and Sexton, K. (2011). Conceptual Models for Cumulative Risk Assessment. *American Journal of Public Health*, 101(S1), S74-S81.

Leventhal, J. (1988). Can child maltreatment be predicted during the perinatal period: Evidence from longitudinal cohort studies? *Reproductive and Infant Psychology* Volume 6, Issue 3, 139-161.

MacMillan, H. L. (2000). Preventive health care, 2000 update: prevention of child maltreatment. *Canadian Medical Association Journal* November 28 2000 163: 1451-1458.

MacMillan, H.L., Wathen, C.N., Barlow, J., Fergusson, D.M., and Leventhal, J.M. (2009). Interventions to prevent child maltreatment and associated impairment. *The Lancet* (Jan 17-Jan 23): 250-66.

Mansell, J., Ota, R., Erasmus, R. and Marks, K. (2011). Reframing child protection: A response to a constant crisis of confidence in child protection, *Children and Youth Services Review*, Volume 33, Issue 11.

Mersky, J.P., Topitzes, J.D. and Reynolds A.J. (2011). Maltreatment prevention through early childhood intervention: A confirmatory evaluation of the Chicago Child–Parent Center preschool program. *Children and Youth Services Review,* 33, 1454–1463.

Ministry of Social Development (2012). *Statistical Report for the year ending 2011*. Wellington: Ministry of Social Development. Available from http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/statistics/statistical-report/statistical-report-2011.html

Moyer, V. A. (2013). Primary care interventions to prevent child maltreatment: U.S. Preventive Service Task Force Recommendation Statement. *Annals of Internal Medecine*. Available from: http://annals.org/article.aspx?articleid=1696071

Munro, E. (2011). *The Munro Review of Child Protection: Final Report, a Child-centred System*. Department for Education, The Stationery Office, London, UK.

Nelson, HD., Selph, S., Bougatsos, C., and Blazina, I. (2013). *Behavioral Interventions and Counseling to Prevent Child Abuse and Neglect: Systematic review to update the U.S. Preventive Services Task Force Recommendation* [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013 Jan. Available from: http://www.ncbi.nlm.nih.gov/books/NBK117232/

New Zealand Government (2012). *White Paper for Vulnerable Children*, Volume II. Available from: http://www.childrensactionplan.govt.nz/the-white-paper/volume-two-of-the-white-paper

Nygren, P., Nelson, HD., and Klein, J. (2004). Screening children for family violence: a review of the Evidence for the U.S. Preventive Services Task Force. *Annals of Family Medicine*. Mar-Apr;2(2):161-9. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1466647/

Panattoni, L.E., Vaithianathan, R., Ashton, T. and Lewis, G.H. (2011). Predictive risk modelling in Health: Options for New Zealand and Australia. *Australian Health Review* 35, 41-51.

Peters, R., and Barlow, J. (2003). Systematic Review of Instruments Designed to Predict Child Maltreatment During the Antenatal and Postnatal Periods. *Child Abuse Review* 12: 416–439.

Putnam-Hornstein, E,. and Needell, B. (2011). Predictors of child welfare contact between birth and age five: an examination of California's 2002 birth cohort. *Children and Youth Services Review*, 33(11), 2400-2407.

Putnam-Hornstein, E. (2012). Preventable injury deaths: a Population-Based Proxy of Child Maltreatment Risk. *Public Health Reports, 127*(2), 163-172.

Putnam-Hornstein, E., Needell, B., and Rhodes, A.E. (2013a). Understanding risk and protective factors for child maltreatment: The value of integrated, population-based data *Child Abuse & Neglect* 37, 116–119.

Putnam-Hornstein, E., Needell, B., King, B. and Johnson-Motoyama, M. (2013b). Racial and ethnic disparities: A population-based examination of risk factors for involvement with child protective services. *Child Abuse & Neglect* 37, 33–46.

Runyan, D., Wattam, C., Ikeda, R., Hassan, F., and Ramiro, L. (2002). Child abuse and neglect by parents and caregivers. In: Krug E, Dahlberg LL, Mercy JA, Zwi AB, Lozano R, editors. *World Report on Violence and Health*. Geneva, Switzerland: World Health Organization.

Schwartz, D., Kaufman, A., and Schwartz, I. (2004). Computational Intelligence Techniques for Risk Assessment and Decision Support. *Children and Youth Services Review*, Volume 26, Issue 11, 1081 – 1095.

Schwartz, I.M., Jones, P., Schwartz, D., and Obradovic, Z. (2008). Improving Social Work through the Use of Technology and Advanced Research Methods, D. Lindsey and A. Shlonsky, *Child Welfare Research.* Oxford University Press.

Sturrock, F. and Gray, D. (2013). *The Incredible Years Pilot Study Evaluation Report*, Ministry of Social Development. Available from http://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/evaluation/incredible-years-report/incredible-years-report.pdf

Tamati, A., Hond-Flavell, E., Korewha, H. and the whānau of Te Kōpae Piripono (2008) *Ko koe kei tēnā kīwai, ko au kei tēnei kīwai o te kete (You carry your handle and I'll carry my handle, of our kete).* Centre of Innovation Research Report of Te Kōpae Piripono Available from: http://www.educationcounts.govt.nz/publications/ECE/22551/34825/34830

Trenholme A., Vogel A., and Lennon D., McBride, C., Stewart, J., Best, E., Mason, H., and Percival, T. (2012). Household characteristics of children under 2 years admitted with lower respiratory tract infection in Counties Manukau, South Auckland. *The New Zealand Medical Journal,* 125(1367).

Vaithianathan, R., Maloney, T., De Haan, I., and Dare, T. (2012). *Vulnerable Children: Can administrative data be used to identify children at risk of adverse outcomes?* University of Auckland, Centre for Applied Research in Economics (CARE). Available from http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/vulnerable-children/index.html

Vaithianathan, R., Maloney, T., Putnam-Hornstein, E. and Jiang, N. (2013). Children in the public benefit system at risk of maltreatment: Identification via predictive modelling. *American Journal of Preventive Medicine,* 45(3), pp 354-359.

Welch, D., and Wilson, M. (2010). *Lifecourse factors associated with time spent receiving benefit in young adulthood: Full report on early findings.* Dunedin Multidisciplinary Health and Development Study, Otago University, and Ministry of Social Development. Available from http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/sole-parenting/index.html

Williams, G. (2011). *Data mining with Rattle and R: The art of excavating data for knowledge discovery.* Springer.

Wilson, J. M. G. and Jungner G. (1968). *Principles and Practice of Screening for Disease.* World Health Organization: Geneva.

Witten, I., Frank E. and Hall M. (2011). *Data mining: practical machine learning tools and techniques* 3rd edition 2011. Morgan Kaufman Publishers.

Wu, S., Ma, C., Carter, R., Ariet, M., Feaver, E., Resnick, M., and Roth, J. (2004). Risk factors for infant maltreatment: a population-based study. *Child Abuse & Neglect*, 28, 1253–1264.

Wulczyn, F. (2009). Epidemiological perspectives on maltreatment prevention. *The Future of Children*, 19 (2), 39-66.