

# Identifying neurodevelopmental disabilities from nationalised preschool health check

Australian & New Zealand Journal of Psychiatry  
2023, Vol. 57(8) 1140–1149  
DOI: 10.1177/00048674231151606

© The Author(s) 2023



Article reuse guidelines:  
sagepub.com/journals-permissions  
journals.sagepub.com/home/anp



Himang Mujoo<sup>1,2</sup>, Nicholas Bowden<sup>1,2</sup>, Hiran Thabrew<sup>1,3</sup>,  
Jesse Kokaua<sup>1,4</sup>, Richard Audas<sup>5</sup> and Barry Taylor<sup>1,2</sup>

## Abstract

**Objective:** Models of psychometric screening to identify individuals with neurodevelopmental disabilities (NDDs) have had limited success. In Aotearoa/New Zealand, routine developmental surveillance of preschool children is undertaken using the Before School Check (B4SC), which includes psychometric and physical health screening instruments. This study aimed to determine whether combining multiple screening measures could improve the prediction of NDDs.

**Methods:** Linked administrative health data were used to identify NDDs, including attention deficit hyperactivity disorder, autism spectrum disorder and intellectual disability, within a multi-year national cohort of children who undertook the B4SC. Cox proportional hazards models, with different combinations of potential predictors, were used to predict onset of a NDD. Harrell's c-statistic for composite models were compared with a model representing recommended cutoff psychometric scores for referral in New Zealand.

**Results:** Data were examined for 287,754 children, and NDDs were identified in 10,953 (3.8%). The best-performing composite model combining the Strengths and Difficulties Questionnaire, the Parental Evaluation of Developmental Status, vision screening and biological sex had 'excellent' predictive power (C-statistic: 0.83) compared with existing referral pathways which had 'poor' predictive power (C-statistic: 0.68). In addition, the composite model was able to improve the sensitivity of NDD diagnosis detection by 13% without any reduction in specificity.

**Conclusions:** Combination of B4SC screening measures using composite modelling could lead to significantly improved identification of preschool children with NDDs when compared with surveillance that rely on individual psychometric test results alone. This may optimise access to academic, personal and family support for children with NDDs.

## Keywords

Paediatrics, neurodevelopment, mental health, screening, childhood, prediction

## Introduction

Data from several countries suggest prevalence of neurodevelopmental disabilities (NDDs) is rising and estimated to affect up to 17% of children (Arora et al., 2018; GBD 2017 Child and Adolescent Health Collaborators et al., 2019; Zablotsky et al., 2019). Improved public awareness, diagnostic practices, surveillance programmes and increasing survival rates from chronic illness are some possible reasons explaining this trend (Arora et al., 2018). NDDs are a group of often co-occurring conditions that include attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), and intellectual disability (ID). Internationally, these are estimated to occur in children and young people at rates of 7.2%, 2.3%, and 1%, respectively

<sup>1</sup>A Better Start National Science Challenge, Liggins Institute, University of Auckland, Auckland, New Zealand

<sup>2</sup>Department of Women's and Children's Health, University of Otago, Dunedin, New Zealand

<sup>3</sup>The Werry Centre, Department of Psychological Medicine, University of Auckland, Auckland, New Zealand

<sup>4</sup>Va'a O Tautai – Centre for Pacific Health, Division of Health Sciences, University of Otago, Dunedin, New Zealand

<sup>5</sup>Faculty of Medicine, Memorial University of Newfoundland and Labrador, St. John's, NL, Canada

### Corresponding authors:

Himang Mujoo, Department of Women's and Children's Health, University of Otago, 201 Great King Street, Dunedin 9016, New Zealand.  
Email: himang.mujoo@health.nsw.gov.au

Nicholas Bowden, Department of Women's and Children's Health, University of Otago, 201 Great King Street, Dunedin 9016, New Zealand.  
Email: nick.bowden@otago.ac.nz

(Maenner et al., 2021; Maulik et al., 2011; Thomas et al., 2015). Long-term health consequences of NDDs include deleterious effects on physical health (Aarons et al., 2008), academic achievement, personal economic outcomes, and social well-being (Capone Singleton, 2018; Duncan et al., 2007). It has been suggested that children with earlier identification of NDDs benefit from interventions during critical periods of their development (Bailey et al., 2001) and provide significant societal savings (Cidav et al., 2017). Therefore, to afford timely support for children with NDDs, several countries, including Aotearoa/New Zealand (NZ), have instituted standardised developmental screening as part of the entry to school (Committee on Children With Disabilities, 2001).

The Before School Check (B4SC) is a nationwide, government-sponsored preschool school health assessment offered to all NZ children, typically at 4 years of age (Ministry of Health, 2016). It is the last in a series of eight free health surveillance checks from birth and reaches over 90% of eligible individuals (Gibb et al., 2019). The B4SC covers child development, behaviour, hearing, vision, oral health, and growth (Statistics New Zealand, 2017b). Currently, children with NDDs may be identified during the B4SC via abnormal scores on either the Strengths and Difficulties Questionnaire (SDQ) or the Parental Evaluation of Developmental Status (PEDS) questionnaire (WHO, 2020; Wilson et al., 1968). Cut scores on individual scales are used to discern potentially indicative behaviour or development and prompt referral for further assessment and support (Radecki et al., 2011). Elevated SDQ scores are associated with poor neurodevelopment and increased risk of NDDs (Becker et al., 2015; Goodman et al., 2003; Nielsen et al., 2019), while total PEDS cutoff scores at the 16th percentile were predictive of developmental delays in children with less than 8 years of age (Brothers et al., 2008). In practice, cut score-dependent identification of individuals with NDDs using the SDQ and PEDS is challenged by non-linear patterns of early development, lack of cross-cultural validation (especially, in the case of NZ Māori and Pacific Island populations), and lack of consensus regarding grounds for onward referral (Mukherjee et al., 2014; Sand et al., 2005). Such issues have also previously been identified when using other developmental screening instruments (Marlow et al., 2019). Thus, there is no ‘gold standard’ screening instrument for NZ children with NDDs, and there are concerns about the under-identification of their prevalence (Kaufman, 2020).

A growing body of evidence shows associations between neurodevelopment and physical health, including biological sex (Martin and Hadwin, 2022), vision (Harris and Lord, 2016), hearing (Istad et al., 2019), oral health (Yusuf et al., 2020) and weight (Wentz et al., 2017). Thus, we hypothesise that combining multiple mental health determinants into a single predictive model will more accurately identify children with NDDs when compared with existing

referral guidelines involving cutoff scores. Examples of composite measures previously used to improve the validity or efficiency of case identification in a wide range of applications include identifying psychiatric disorders (Kessler et al., 1998; Liu et al., 2021), the onset of dementia (Borson et al., 2000), and identifying problematic drinking in students (Werner et al., 1994). In clinical research, motor, visual and language development measures have been previously combined to predict cognitive functioning (Shank, 2011).

In NZ, the Integrated Data Infrastructure (IDI), a large database containing anonymised microdata, including on the health of children, is available for research deemed to be for the public good. Although the definitions of ‘best practice’ in the use of such large administrative data are currently evolving (Grath-Lone et al., 2022; Harron et al., 2017), it nonetheless provides opportunities to explore health associations. In this study, we explore novel uses of IDI to improve prediction and referral pathways for children with NDDs. In the absence of any national registries for NDDs in NZ, previously established IDI-based methods will be used to identify cases of NDDs using a combination of diagnostic data and pharmaceutical dispensing (Bowden et al., 2020a, 2020b). These methods remain formally unvalidated; however, they have been applied in extant literature with demonstrated face validity (Bowden et al., 2020c, 2022; McLay et al., 2021, 2022). Physical and psychometric health surveillance measures contained in the B4SC data set, together with demographic attributes, will be combined with the aim to

1. Apply predictive models and ascertain whether it would be possible to increase the accuracy of the health assessment in identifying children with NDDs,
2. Identify the combination of predictors that optimise the performance of the model, and
3. Offer recommendations for improving B4SC-based identification of children with NDDs.

## Methods

### Study design

This was a retrospective, time-to-event analysis of a multi-year national cohort of NZ preschool children. The study has ethical approval under the minimal risk category by the University of Otago Human Research Ethics Committee (Approval: HD17/004).

### Data source: the integrated data infrastructure

Data for the study were sourced from the Statistics NZ-curated Integrated Data Infrastructure (IDI), a national repository of administrative and survey data that includes

**Table 1.** IDI codes used to identify NDDs.

|      | DSM-IV (PRIMHD)               | ICD-10-AM (PRIMHD & NMDS)                | Pharmaceuticals (PHARMS)   | Assigned diagnoses (Socrates)   |
|------|-------------------------------|--|--|---|
| ADHD | 314.00, 314.01, 314.9         | F90.0, F90.8, F90.9                      | 1389 (dexamphetamine), 1809 (methylphenidate), 3880 (methylphenidate extended-release), 3887 (atomoxetine) | 1201 (ADHD)   |
| ASD  | 299.00, 299.10, 299.80        | F84.0, F84.1, F84.3, F84.5, F84.8, F84.9 | n/a  | 1211 (ASD), 1206 (Asperger's syndrome), 1207 (other ASD)  |
| ID   | 317, 319, 318.0, 318.1, 318.2 | F70–F79                                  | n/a  | 1208 (intellectual disability), 1209 (learning disability), 1210 (developmental delay), 1299 (other intellectual, learning or developmental disorder) |

DSM-IV: *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition; PRIMHD: Programme for the Integration of Mental Health Data; ICD-10-AM: The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification; NMDS: National Minimum Dataset; PHARMS: Pharmaceutical Collection; ADHD: attention deficit hyperactivity disorder; ASD: autism spectrum disorder; ID: intellectual disability.

B4SC screening data (Statistics New Zealand, 2017a). De-identified, probabilistically linked IDI data are available to trusted researchers operating within the Five Safes framework (Statistics New Zealand, 2017a). Access to IDI data was approved by Statistics NZ.

### Participants

The cohort identified for this study included NZ preschool (4–5 years old) children who had available B4SC data within the IDI, between 1 January 2011 and 30 June 2016 (Gibb et al., 2016; Zhao et al., 2018). Age was determined at the time of B4SC completion. Individuals were excluded from the analysis if their age was beyond the B4SC age bracket (<4 or >5 years), where their B4SC was incomplete, or if consent was not provided.

### Primary outcomes: neurodevelopmental disorders

The NDDs of interest, ADHD, ASD, ID, and a combined category termed 'any NDD' were identified using established case identification methods for IDI-based data (Bowden et al., 2020a, 2020b). These methods utilise diagnostic codes and pharmaceutical dispensing information obtained from four health data sets: The Programme for the Integration of Mental Health Data (PRIMHD), a national collection of publicly funded specialist mental health service use data; the National Minimum Dataset (NMDS), a national collection of publicly funded hospital admissions; Socrates, a national database of the Ministry of Health's Disability Support Services clients; and the Pharmaceutical Collection (PHARMS), a national claims and payment information data set from pharmacists for government-subsidised medication dispensing. Each NDD was identified

using up to 9 years of data from 1 July 2010 until 30 June 2019 via the presence of one or more diagnostic codes across any of the four data sets (see Table 1 for details). Dichotomous indicators (yes/no) were constructed for each specific NDD and for a composite indicator termed 'any NDD'.

### Covariates

Predictors of NDDs included developmental, behavioural, and physical health measures from the B4SC between 1 January 2010 (the point at which the B4SC is considered universal (Gibb et al., 2019) until 30 June 2016 (to enable at least 3 years of follow-up for each child in the cohort). Detailed methods for capturing these measures can be found in *The B4 School Check Handbook* (Ministry of Health, 2008). Biological sex was also included in the models.

*Strengths and Difficulty Questionnaire (SDQ)*. The SDQ encompasses 25 questions, reported by parents and teachers, which identify five psychosocial attributes, namely conduct, emotional, hyperactivity, peer problems, and pro-social behaviour (Goodman, 1997; Goodman et al., 2000). In the present study, scores of individual subscale measures and a summary score representing overall difficulty (SDQ Total) were calculated from the first four sub-scales. Per the literature, SDQ scores were excluded where all five subscales registered as zero (Kersten et al., 2018; Ministry of Health, 2015; Silva et al., 2015).

Approximately half of the cohort did not receive an SDQ assessment from their teachers. To make the best use of available teacher-assessed SDQ scores, parent and teacher scores were normalised and missing teacher scores were imputed with normalised parent scores where available.

Two additional binary indicators were created for subsequent referral based on recommended governmental guidelines, SDQ Total scores:  $\geq 17$  from a parent-completed questionnaire; and  $\geq 16$  from an imputed teacher-completed questionnaire.

**Parents' Evaluation of Developmental Status (PEDS).** PEDS examines parental concerns regarding development (cognition, communication and motor skills) in children up to 8 years of age (Glascoe, 2003; The Royal Children's Hospital Melbourne, 2108). PEDS includes 10 questions about behaviour; development; speech and language; and fine or gross motor skills. Two continuous scores were generated from PEDS responses and used in the present study: (1) PEDS with significant concerns (PEDS-sig), which captures established predictors of developmental delay; and (2) PEDS other concerns (PEDS-ns), while not considered established predictors of developmental delay may still highlight issues that require ongoing monitoring. An additional PEDS measure, PEDS-sig referral was constructed to represent governmental guidelines for a referral if two or more significant concerns are reported.

**Hearing, vision, and oral test.** Hearing, vision, and oral test results were captured by specialist services during B4SC screening. Each measure used in this study was converted into a dichotomous variable, with '1' indicating a pass of the respective check and '0' indicating referral to an appropriate service when concerns were raised. The auditory assessment included sweep audiometry with follow-up tympanometry test if audiometry were equivocal or abnormal. Vision was assessed using Snellen charts. Oral health screening involved the 'Lift the Lip' check to identify dental decay (Arrow et al., 2013).

**Height and Weight measurements.** As per guidelines, anthropometric measurements were conducted using calibrated instruments (Ministry of Health, 2008). Height ( $\pm 0.1$  cm) was measured using a portable stadiometer (Leicester Height Measure or SECA 214) and weight ( $\pm 0.1$  kg) was measured using a floor scale (SECA 862, SECA 770, Tanita WB 100S MA or Tanita HD-351). The WHO Anthro Macro (version 3.2.2) was used to obtain sex-specific body mass index (BMI)-for-age Z-scores (WHO Anthro Department of Nutrition, 2019). A three-level categorical variable, BMI, was used to indicate children within a healthy BMI range of plus or minus one standard deviation (*SD*) of the WHO mean for children at 4 years of age, those one *SD* below the mean considered underweight, and one *SD* above the mean considered overweight (de Onis et al., 2019).

**Biological sex.** Information on the biological sex of participants was sourced from the personal details table in the IDI and categorised as either female or male.

## Procedure

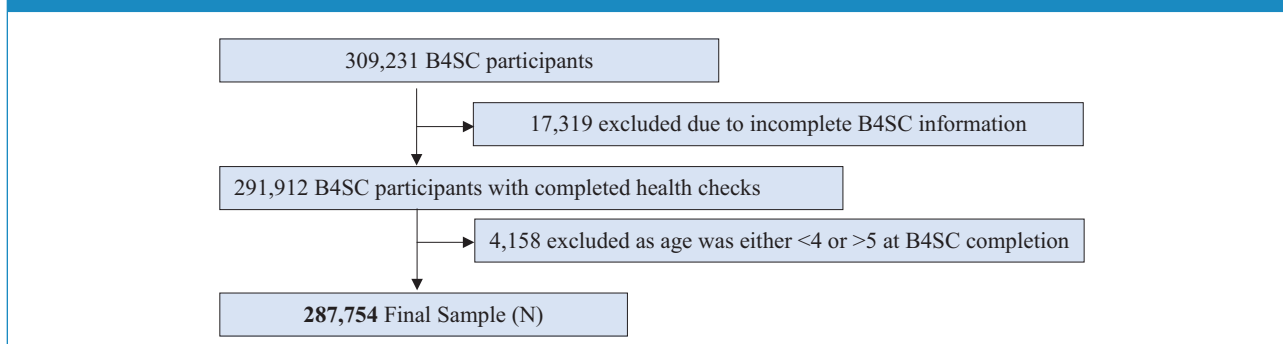
Data were accessed from the June 2020 refresh of the IDI and extracted using SAS version 7.1 (SAS Institute Inc, 2014). The resulting data were analysed using StataMP version 15 (StataCorp, 2017). Following Statistics NZ confidentiality requirements, all counts were randomly rounded to base 3 and suppressed if less than six to minimise any risk of individual identification (Statistics New Zealand, 2019). Confidentialised results were approved for final release by Statistics NZ.

## Statistical analysis

Observed rates of NDDs (any NDD, ASD, ADHD, and ID) were presented overall and stratified by sex. Unadjusted hazard ratios (HRs) and associated 95% confidence intervals (CIs) for each predictor measure on NDD status were generated from Cox proportional hazards regression models (StataCorp, 2019). Onset for the Cox models of NDD was registered as the earliest time to (1) first NDD diagnosis, (2) death, or (3) a period of overseas travel greater than 3 months. Otherwise, times were censored for individuals that survived to the end of the study period (30 June 2019). The proportional hazards assumption was checked for all predictors using Schoenfeld residual plots (Schoenfeld, 1980).

For each NDD classification, baseline models (one) were constructed utilising current referral pathway information from SDQ and PEDS scores at recommended thresholds for a referral (Ministry of Health, 2008). Baseline models were compared with a sequence of Cox models which incrementally combined predictor variables: M1 (SDQ subscales as continuous variables), M2 (PEDS scores), M3 (M1 + M2), M4 (M3 + Hearing), M5 (M3 + Vision), M6 (M3 + Oral), M7 (M3 + BMI) and M8 (M5 + Sex).

The predictive ability of each model was assessed using Harrel's c-statistic, a performance measure that assesses models of dichotomous outcomes (Pencina et al., 2008). Using the c-statistic, the discriminatory power of models is categorised as 'poor' ( $\leq 0.7$ ), 'acceptable' (0.7–0.79), 'excellent' (0.80–0.89), or 'outstanding' ( $\geq 0.9$ ) (Hosmer and Lemeshow, 2000). Models were assessed further using positive predictive value (PPV), negative predictive value (NPV), sensitivity, and specificity. For the baseline model, actual referrals were compared with observed NDD diagnoses. For each subsequent model, predicted probabilities for NDDs were calculated and compared with actual observed NDDs. The threshold value for predicted probabilities was set to constrain the percentage of the referred population to that of the baseline model (9.3%). Thresholds were adjusted to halve (4.7%) and double referrals (18.6%) to reflect differing capacities of public health funding.

**Figure 1.** Participant flow chart.

## Results

### Participant population

Over the study period, a total of 309,231 children were administered a B4SC, of which 287,754 children were aged between 4 to 5 years and had complete B4SC information and consent, which were therefore eligible for inclusion in our analysis (see Figure 1).

### Participants with identifiable NDDs

NDDs were identified in 10,953 children (3.8% of the participant population) (Table 2). ADHD was the most common (2.5%), followed by ASD (1.3%) and ID (1.0%). Overall, boys with NDDs outnumbered girls by over three to one. This sex difference was particularly high for ADHD and ASD at approximately four to one.

### Prediction of NDDs by B4SC components

Among the analytical sample, completion rates for each B4SC component were typically high, including SDQ parent (98.5%), PEDS (99.1%), vision (100.0%), hearing (92.3), oral (99.1%), and height and weight (99.1%). The

SDQ teacher was administered at a far lower rate (51.4%). For any of the NDDs investigated, unadjusted HRs for SDQ and PEDS referral thresholds were significantly greater than one, indicating a higher risk of any NDD against those who did not meet the referral threshold in each score (Table 3). Likewise, HRs for all continuous SDQ scores (except SDQ Pro-social), both PEDS scores, high BMI, and male sex were associated with significantly higher risk of any NDD. In contrast, both SDQ Pro-social scores, a pass on vision testing, hearing testing, and normal oral health were associated with significantly lower risk of any NDD. Having a low BMI (one *SD* below the mean) had no association with any NDD ( $p$  value=0.58). Similar patterns were replicated within each individual NDD diagnoses. The only exception was children with a BMI one *SD* below the mean, who were at no more risk of individual NDDs than children with normal BMI, except in the case for ID which was associated with increased risk.

### Comparing the predictive performance of models

Baseline models had 'poor' discriminatory power for any NDD (c-statistic=0.68) and ADHD (0.66) but had 'acceptable' power for ASD (0.72) and ID (0.78). In contrast, the

**Table 2.** NDDs (ADHD, ASD, ID) identified and population rates among the eligible sample of 287,754 children.

|         | Total<br><i>n</i> (%) | Male<br><i>n</i> (%) | Female<br><i>n</i> (%) |
|---------|-----------------------|----------------------|------------------------|
| Any NDD | 10,953 (3.8)          | 8,508 (5.8)          | 2,445 (1.7)            |
| ADHD    | 7,185 (2.5)           | 5,745 (3.9)          | 1,437 (1.0)            |
| ASD     | 3,597 (1.3)           | 2,910 (2.0)          | 690 (0.5)              |
| ID      | 2,820 (1.0)           | 2,004 (1.4)          | 816 (0.6)              |
| No NDD  | 276,804 (96.2)        | 139,338 (94.2)       | 137,463 (98.3)         |
| Total   | 287,754               | 147,849              | 139,908                |

NDD: neurodevelopmental disabilities; ADHD: attention deficit hyperactivity disorder; ASD: autism spectrum disorder; ID: intellectual disability.

**Table 3.** Unadjusted hazard ratios (HRs) and associated 95% CIs for B4SC measures and sex on NDDs.

| Predictive factors                 | N       | NDD HR (95% CI)  | ADHD HR (95% CI) | ASD HR (95% CI)     | ID HR (95% CI)      |
|------------------------------------|---------|------------------|------------------|---------------------|---------------------|
| SDQ referral (Parent: $\geq 17$ )  | 283,431 | 6.68 (6.34–7.04) | 6.36 (5.99–6.75) | 8.12 (7.38–8.92)    | 7.46 (6.69–8.32)    |
| SDQ referral (Teacher: $\geq 16$ ) | 147,849 | 8.38 (7.80–9.00) | 7.43 (6.85–8.07) | 12.12 (10.64–13.80) | 11.26 (9.61–13.21)  |
| PEDS-sig referral ( $\geq 2$ )     | 285,261 | 5.64 (5.38–5.92) | 3.99 (3.76–4.23) | 9.07 (8.35–9.85)    | 13.78 (12.53–15.15) |
| SDQ Total (Parent)                 | 283,431 | 1.16 (1.16–1.17) | 1.16 (1.15–1.16) | 1.17 (1.17–1.18)    | 1.17 (1.17–1.18)    |
| SDQ Total (Teacher)                | 147,849 | 1.16 (1.16–1.16) | 1.15 (1.15–1.16) | 1.18 (1.17–1.19)    | 1.18 (1.17–1.19)    |
| SDQ Pro-social (Parent)            | 283,437 | 0.77 (0.76–0.77) | 0.80 (0.79–0.81) | 0.69 (0.68–0.70)    | 0.73 (0.72–0.75)    |
| SDQ Pro-social (Teacher)           | 147,858 | 0.74 (0.74–0.75) | 0.77 (0.76–0.77) | 0.67 (0.66–0.69)    | 0.70 (0.68–0.72)    |
| SDQ Conduct (Parent)               | 283,437 | 1.41 (1.40–1.43) | 1.44 (1.43–1.46) | 1.38 (1.35–1.40)    | 1.36 (1.34–1.39)    |
| SDQ Conduct (Teacher)              | 147,855 | 1.40 (1.38–1.41) | 1.41 (1.39–1.43) | 1.39 (1.36–1.42)    | 1.35 (1.31–1.38)    |
| SDQ Emotional (Parent)             | 283,431 | 1.20 (1.19–1.21) | 1.15 (1.13–1.16) | 1.32 (1.30–1.35)    | 1.32 (1.28–1.35)    |
| SDQ Emotional (Teacher)            | 147,849 | 1.17 (1.15–1.19) | 1.10 (1.08–1.13) | 1.34 (1.31–1.38)    | 1.30 (1.25–1.34)    |
| SDQ Hyperactivity (Parent)         | 283,437 | 1.44 (1.43–1.46) | 1.46 (1.45–1.48) | 1.37 (1.35–1.39)    | 1.48 (1.46–1.51)    |
| SDQ Hyperactivity (Teacher)        | 147,858 | 1.41 (1.40–1.43) | 1.41 (1.40–1.42) | 1.40 (1.38–1.43)    | 1.52 (1.49–1.55)    |
| SDQ Peer Problems (Parent)         | 283,437 | 1.42 (1.41–1.44) | 1.32 (1.30–1.33) | 1.69 (1.66–1.71)    | 1.55 (1.52–1.59)    |
| SDQ Peer Problems (Teacher)        | 147,858 | 1.41 (1.40–1.43) | 1.41 (1.40–1.42) | 1.40 (1.38–1.43)    | 1.52 (1.49–1.55)    |
| PEDS-sig                           | 285,258 | 1.91 (1.88–1.93) | 1.66 (1.63–1.69) | 2.11 (2.06–2.16)    | 2.18 (2.13–2.22)    |
| PEDS-ns                            | 285,258 | 1.54 (1.53–1.55) | 1.49 (1.47–1.51) | 1.63 (1.60–1.65)    | 1.66 (1.63–1.69)    |
| Vision (not referred)              | 287,694 | 0.42 (0.40–0.44) | 0.51 (0.49–0.54) | 0.40 (0.36–0.43)    | 0.15 (0.14–0.17)    |
| Hearing (not referred)             | 265,602 | 0.63 (0.59–0.67) | 0.75 (0.70–0.81) | 0.51 (0.46–0.57)    | 0.28 (0.25–0.32)    |
| Oral (not referred)                | 285,051 | 0.85 (0.80–0.91) | 0.87 (0.81–0.94) | 0.90 (0.79–1.02)    | 0.75 (0.66–0.85)    |
| WHO BMI–overweight risk            | 285,216 | 1.19 (1.14–1.24) | 1.15 (1.10–1.21) | 1.31 (1.20–1.42)    | 1.34 (1.21–1.47)    |
| WHO BMI–underweight risk           | 285,219 | 1.03 (0.93–1.14) | 0.97 (0.85–1.10) | 0.93 (0.76–1.14)    | 1.36 (1.11–1.67)    |
| Sex (male)                         | 287,754 | 3.49 (3.32–3.66) | 3.84 (3.62–4.07) | 4.05 (3.67–4.48)    | 2.33 (2.11–2.58)    |

NDD: neurodevelopmental disabilities; HR: hazard ratios; CI: confidence interval; ADHD: attention deficit hyperactivity disorder; ASD: autism spectrum disorder; ID: intellectual disability; SDQ: Strengths and Difficulties Questionnaire; WHO: World Health Organization; BMI: body mass index.

composite models consistently outperformed the Baseline models (see supplementary material for complete regression results).

Combining all five SDQ sub-scales and the PED scores into a predictive model yielded higher c-statistics (M3) when compared with baseline models or individually modelled measures (M1, M2), thus indicating improved predictive power (Table 4). The c-statistic further improved across all NDD classifications with the addition of ocular health (M5) and biological sex (M8), but not auditory health (M4), oral health (M6), or BMI (M7). Ultimately, model M8

outperformed other models with ‘excellent’ discriminatory power for any NDD (0.83), ADHD (0.83), ASD (0.86), and ‘outstanding’ for ID (0.90).

These results were further confirmed by comparing the PPV, NPV, sensitivity, and specificity values of the baseline and M8 the best-performing composite model (Table 5). The composite models performed better than or equal to the baseline for all metrics for all NDD classifications. For any NDD, sensitivity improved by 13% (46% vs 52%). Improvement in sensitivity was highest for ADHD (38% vs 50%). These gains in sensitivity were observed without loss

**Table 4.** Predictive assessment (C-stats<sup>a</sup>) of regression models.

| Model                 | Included variables | NDD<br>HR (95% CI) | ADHD<br>HR (95% CI) | ASD<br>HR (95% CI) | ID<br>HR (95% CI) |
|-----------------------|--------------------|--------------------|---------------------|--------------------|-------------------|
| Baseline <sup>b</sup> | SDQ-PEDS referral  | 0.68 (0.67–0.68)   | 0.66 (0.65–0.67)    | 0.72 (0.71–0.74)   | 0.78 (0.76–0.79)  |
| M1                    | SDQ Subscales      | 0.80 (0.79–0.81)   | 0.80 (0.80–0.81)    | 0.82 (0.81–0.84)   | 0.84 (0.83–0.86)  |
| M2                    | PEDS sig & ns      | 0.74 (0.73–0.74)   | 0.71 (0.70–0.72)    | 0.80 (0.78–0.81)   | 0.82 (0.81–0.84)  |
| M3                    | M1 + M2            | 0.82 (0.81–0.82)   | 0.81 (0.81–0.82)    | 0.85 (0.83–0.86)   | 0.88 (0.87–0.89)  |
| M4                    | M3 + Hearing       | 0.81 (0.80–0.82)   | 0.81 (0.80–0.82)    | 0.83 (0.81–0.84)   | 0.87 (0.85–0.88)  |
| M5                    | M3 + Vision        | 0.82 (0.81–0.83)   | 0.81 (0.81–0.82)    | 0.85 (0.84–0.86)   | 0.90 (0.89–0.91)  |
| M6                    | M3 + Oral          | 0.82 (0.81–0.82)   | 0.81 (0.80–0.82)    | 0.85 (0.83–0.86)   | 0.88 (0.87–0.89)  |
| M7                    | M3 + BMI           | 0.82 (0.81–0.82)   | 0.81 (0.80–0.82)    | 0.84 (0.83–0.86)   | 0.88 (0.87–0.89)  |
| M8                    | M5 + Sex           | 0.83 (0.83–0.84)   | 0.83 (0.82–0.84)    | 0.86 (0.85–0.88)   | 0.90 (0.89–0.91)  |

NDD: neurodevelopmental disabilities; HR: hazard ratios; CI: confidence interval; ADHD: attention deficit hyperactivity disorder; ASD: autism spectrum disorder; ID: intellectual disability; PEDS: Parental Evaluation of Developmental Status; SDQ: Strengths and Difficulties Questionnaire; BMI: body mass index.

<sup>a</sup>Discriminatory power of models: poor discrimination ( $\leq 0.7$ ), acceptable (0.7–0.79), excellent (0.80–0.89), or outstanding ( $\geq 0.9$ ) (Hosmer and Lemeshow, 2000).

<sup>b</sup>The baseline C-stats were derived following modelling of referral guidelines utilising PEDS and total SDQ scores.

of specificity. For individual NDDs, doubling the referral rate improved sensitivity from baseline but specificity declined. In contrast, specificity improved by halving referrals at the expense of sensitivity. Those effects were also reflected with varying magnitude in NPV and PPV.

## Discussion

### Key findings

In this national retrospective cohort study, we sought to combine psychometric and physical health screening measures into a novel composite predictive model to improve the identification of NDDs. Composite modelling was found to improve the identification of preschool children with NDDs compared with existing referral pathways. The optimal composite model, including indicators of SDQ scales, PEDS, vision, and biological sex, had excellent discriminatory power (C-statistic 0.83) (Hosmer and Lemeshow, 2000). It outperformed all other models including the current use of cutoff scores from individual psychometric questionnaires. Using this model, the sensitivity of NDD detection was improved by 13% without loss of specificity.

Factors that may limit further clinical support following B4SC assessment include barriers to accessing local resources; some children are already known to child health and development providers; and lack of parental resources to engage with further assessment (Richards et al., 2019). Improving the sensitivity of NZ's current system of health checks could have a role in addressing some of these issues

as the higher sensitivity will improve the cost-efficiency of the whole system. Given that the B4SC is broadly accepted in NZ (Williams, 2013), it would be advantageous to retain current screening tools such as the SDQ and PEDS while improving their effectiveness by employing prediction models that combine health screening measures. Such models should be further validated among specific communities with different levels of need.

### Strengths and limitations

Strengths of this study include using a large, national data set and well-established statistical modelling techniques. The approach is novel concerning the application to NDDs. However, the findings must be viewed considering several limitations. First, cases of NDDs are established using unvalidated methods that likely undercount true NDD prevalence, and the extent of false positives remains unknown (Bowden et al., 2020a, 2020b). Second, in practice, the process of referral often involves communication between professionals, which is not available in the IDI and consequently not captured in the baseline model. Third, B4SC administrative data were not originally intended for health analysis and therefore has issues surrounding the reliability of individual B4SC measures (Schluter et al., 2018). Finally, our approach of combining teacher and parent SDQ measures to account for the low rate of teacher administered SDQs completed is imperfect, particularly given parent and teacher SDQ scores do not correlate well (Kersten et al., 2018). Future work should look to explore alternative imputation methods, and

**Table 5.** Comparisons of PPV, NPV, sensitivity and specificity between baseline and prediction model.

| Model                  | NDD (%) |     | ADHD (%) |     | ASD (%) |      | ID (%) |      |
|------------------------|---------|-----|----------|-----|---------|------|--------|------|
|                        | PPV     | NPV | PPV      | NPV | PPV     | NPV  | PPV    | NPV  |
| Baseline <sup>a</sup>  | 18%     | 98% | 10%      | 98% | 7%      | 99%  | 6%     | 100% |
| M8                     | 21%     | 98% | 13%      | 99% | 8%      | 99%  | 7%     | 100% |
| M8-half <sup>b</sup>   | 31%     | 98% | 19%      | 98% | 14%     | 99%  | 12%    | 100% |
| M8-double <sup>c</sup> | 13%     | 98% | 9%       | 99% | 5%      | 100% | 4%     | 100% |

NDD: neurodevelopmental disabilities; ADHD: attention deficit hyperactivity disorder; ASD: autism spectrum disorder; ID: intellectual disability; PPV: positive predictive value; NPV: negative predictive value.

<sup>a</sup>The baseline statistics for each disorder group were derived by comparing actual referrals with observed.

<sup>b</sup>Threshold value set such that in comparison with the baseline model, only half the number of people would be referred.

<sup>c</sup>Threshold value set such that in comparison with the baseline model, twice the number of people would be referred.

ways to improve SDQ teacher completion rates such as protected teacher time or alternative incentives.

Ethical issues associated with the use of IDI data for research include risks of marginalisation, mislabelling, and exacerbation of the apparent boundaries between health and ill health. For a more detailed discussion, see Bowden et al. (2019). Use of IDI data may be disadvantageous for ethnic groups that receive poorly equitable services, such as Māori and Pasifika (Wikaire et al., 2016; Williams et al., 2018). In particular, B4SC participation has been found to be lower for Māori, Pasifika, and those living in socioeconomically deprived areas (Gibb et al., 2019).

### Future directions

Further research is needed to explore the psychometric properties of the composite screening model (M8) with children of different ethnicities, especially Māori and Pasifika. Further addition of predictors (e.g. known co-occurring conditions) could improve the effectiveness of the composite model. Finally, a trial of the composite screening model with a clinical sample would validate its impact on onward referral in the real world.

### Conclusion

Our composite prediction models demonstrated better discrimination for NDDs (ADHD, ASD, and ID) than the current screening guidelines. Thus, we propose that by routinely combining measures of SDQ, PEDS, visual health, and biological sex, earlier detection of children with NDDs and subsequently better support will be possible.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding


The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was part of A Better Start National Science Challenge, funded by the New Zealand Ministry of Business, Innovation and Employment (MBIE) grants UOAX1511 and UOAX1901. The authors would like to further acknowledge Dean’s Medical Student Research Scholarship. The funder/sponsor did not participate in the work.

### Disclaimer

These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) which is carefully managed by Stats NZ. For more information about the IDI, please visit <https://www.stats.govt.nz/integrated-data/>.



## ORCID iDs

Himang Mujoo  <https://orcid.org/0000-0001-7681-0534>

Nicholas Bowden  <https://orcid.org/0000-0003-4589-9956>

Hiran Thabrew  <https://orcid.org/0000-0002-8755-6217>

## Supplemental material

Supplemental material for this article is available online.

## References

- Aarons GA, Monn AR, Leslie LK, et al. (2008) Association between mental and physical health problems in high-risk adolescents: A longitudinal study. *The Journal of Adolescent Health* 43: 260–267.
- Arora NK, Nair MKC, Gulati S, et al. (2018) Neurodevelopmental disorders in children aged 2–9 years: Population-based burden estimates across five regions in India. *PLOS Medicine* 15: e1002615.
- Arrow P, Raheb J and Miller M (2013) Brief oral health promotion intervention among parents of young children to reduce early childhood dental decay. *BMC Public Health* 13: 245–245.
- Bailey JRDB, Bruer JT, Symons FJ, et al. (2001) *Critical Thinking about Critical Periods*. Baltimore, MD: Paul H Brookes Publishing.
- Becker A, Rothenberger A, Sohn A, et al. (2015) Six years ahead: A longitudinal analysis regarding course and predictive value of the Strengths and Difficulties Questionnaire (SDQ) in children and adolescents. *European Child & Adolescent Psychiatry* 24: 715–725.
- Borson S, Scanlan J, Brush M, et al. (2000) The Mini-Cog: A cognitive ‘vital signs’ measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry* 15: 1021–1027.
- Bowden N, Gibb S, Audas R, et al. (2022) Association between high-need education-based funding and school suspension rates for Autistic students in New Zealand. *JAMA Pediatrics* 176: 664–671.
- Bowden N, Gibb S, Thabrew H, et al. (2019) IDI trends in antidepressant dispensing to New Zealand children and young people between 2007/08 and 2015/16. *The New Zealand Medical Journal* 132: 48–61.
- Bowden N, Gibb S, Thabrew H, et al. (2020a) Case identification of mental health and related problems in children and young people using the New Zealand Integrated Data Infrastructure. *BMC Medical Informatics and Decision Making* 20: 42.
- Bowden N, Thabrew H, Kokaua J, et al. (2020b) Autism spectrum disorder/Takiwātanga: An integrated data infrastructure-based approach to autism spectrum disorder research in New Zealand. *Autism* 24: 2213–2227.
- Bowden N, Thabrew H, Kokaua J, et al. (2020c) National prescribing rates and polypharmacy for children and young people in New Zealand with and without autism spectrum disorder. *Research in Autism Spectrum Disorders* 78: 101642.
- Brothers KB, Glascoe FP and Robertshaw NS (2008) PEDS: Developmental milestones – an accurate brief tool for surveillance and screening. *Clinical Pediatrics* 47: 271–279.
- Capone Singleton N (2018) Late talkers: Why the wait-and-see approach is outdated. *Pediatric Clinics of North America* 65: 13–29.
- GBD 2017 Child and Adolescent Health Collaborators, Reiner RC Jr, Olsen HE, et al. (2019) Diseases, injuries, and risk factors in child and adolescent health, 1990 to 2017: Findings from the global burden of diseases, injuries, and risk factors 2017 study. *JAMA Pediatrics* 173: e190337.
- Cidav Z, Munson J, Estes A, et al. (2017) Cost offset associated with early start Denver model for children with Autism. *Journal of the American Academy of Child and Adolescent Psychiatry* 56: 777–783.
- Committee on Children With Disabilities (2001) Developmental surveillance and screening of infants and young children. *Pediatrics* 108: 192–195.
- de Onis M, Borghi E, Arimond M, et al. (2019) Prevalence thresholds for wasting, overweight and stunting in children under 5 years. *Public Health Nutrition* 22: 175–179.
- Duncan GJ, Dowsett CJ, Claessens A, et al. (2007) School readiness and later achievement. *Developmental Psychology* 43: 1428–1446.
- Gibb S, Bycroft C and Matheson-Dunning N (2016) *Identifying the New Zealand Resident Population in the Integrated Data Infrastructure (IDI)*. Wellington, New Zealand: Statistics New Zealand.
- Gibb S, Milne B, Shackleton N, et al. (2019) How universal are universal preschool health checks? An observational study using routine data from New Zealand’s B4 School Check. *BMJ Open* 9: e025535.
- Glascoe FP (2003) Parents’ evaluation of developmental status: How well do parents’ concerns identify children with behavioral and emotional problems? *Clinical Pediatrics* 42(2): 133–138.
- Goodman R (1997) The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 38: 581–586.
- Goodman R, Ford T, Simmons H, et al. (2003) Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *International Review of Psychiatry* 15: 166–172.
- Goodman R, Renfrew D and Mullick M (2000) Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European Child & Adolescent Psychiatry* 9: 129–134.
- Grath-Lone LM, Jay MA, Blackburn R, et al. (2022) What makes administrative data ‘research-ready’? A systematic review and thematic analysis of published literature. *International Journal of Population Data Science* 7: 1718.
- Harris J and Lord C (2016) Mental health of children with vision impairment at 11 years of age. *Developmental Medicine and Child Neurology* 58: 774–779.
- Harron K, Dibben C, Boyd J, et al. (2017) Challenges in administrative data linkage for research. *Big Data & Society* 4: 2053951717745678.
- Hosmer DW and Lemeshow S (2000) Introduction to the logistic regression model. In: *Applied Logistic Regression, Second Edition*. New York: Wiley–Blackwell, pp. 1–30.
- Idstad M, Tams K, Aarhus L, et al. (2019) Childhood sensorineural hearing loss and adult mental health up to 43 years later: Results from the HUNT study. *BMC Public Health* 19: 168.
- Kaufman NK (2020) Rethinking ‘gold standards’ and ‘best practices’ in the assessment of autism. *Applied Neuropsychol Child* 11: 529–540.
- Kersten P, Vandal AC, Elder H, et al. (2018) Strengths and Difficulties Questionnaire: Internal validity and reliability for New Zealand preschoolers. *BMJ Open* 8: e021551.
- Kessler RC, Andrews G, Mroczek D, et al. (1998) The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF). *International Journal of Methods in Psychiatric Research* 7: 171–185.
- Liu YS, Chokka S, Cao B, et al. (2021) Screening for bipolar disorder in a tertiary mental health centre using EarlyDetect: A machine learning-based pilot study. *Journal of Affective Disorders Reports* 6: 100215.
- Maenner MJ, Shaw KA, Bakian AV, et al. (2021) Prevalence and characteristics of Autism Spectrum Disorder among children aged 8 years – Autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveillance Summaries* 70: 1–16.
- Marlow M, Servili C and Tomlinson M (2019) A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: Recommendations for use in low- and middle-income countries. *Autism Research* 12: 176–199.
- Martin J and Hadwin JA (2022) The roles of sex and gender in child and adolescent mental health. *JCPP Advances* 2: e12059.
- Maulik PK, Mascarenhas MN, Mathers CD, et al. (2011) Prevalence of intellectual disability: A meta-analysis of population-based studies. *Research in Developmental Disabilities* 32: 419–436.

- McLay LK, Bowden NJ, Eggleston MJF, et al. (2022) Melatonin dispensing and polypharmacy rates for New Zealand children with autism or attention deficit hyperactivity disorders: A nationwide pharmacoepidemiological study. *Research in Autism Spectrum Disorders* 93: 101948.
- McLay LK, Schluter PJ, Eggleston MJF, et al. (2021) Melatonin dispensing among New Zealand children aged 0–18 years with autism: A nationwide cross-sectional study. *Sleep Medicine* 80: 184–192.
- Ministry of Health (2008) *The B4 School Check: A Handbook for Practitioners*. Wellington, New Zealand: Ministry of Health.
- Ministry of Health (2015) A validation and norming study of the Strengths and Difficulties Questionnaire in the New Zealand context. Available at: <https://www.health.govt.nz/publication/validation-and-norming-study-strengths-and-difficulties-questionnaire-new-zealand-context> (accessed 22 April 2020).
- Ministry of Health (2016) B4 school check information for the health sector. Available at: <https://www.health.govt.nz/our-work/life-stages/child-health/b4-school-check/b4-school-check-information-health-sector> (accessed 31 January 2019).
- Mukherjee SB, Aneja S, Krishnamurthy V, et al. (2014) Incorporating developmental screening and surveillance of young children in office practice. *Indian Pediatrics* 51: 627–635.
- Nielsen LG, Rimvall MK, Clemmensen L, et al. (2019) The predictive validity of the Strengths and Difficulties Questionnaire in preschool age to identify mental disorders in preadolescence. *PLoS ONE* 14: e0217707.
- Pencina MJ, D'agostino RB and Vasan RS (2008) Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Statistics in Medicine* 27: 157–172.
- Radecki L, Sand-Loud N, O'Connor KG, et al. (2011) Trends in the use of standardized tools for developmental screening in early childhood: 2002–2009. *Pediatrics* 128: 14–19.
- Richards N, Reith D, Stitely M, et al. (2019) The Before School Check (B4SC): Reporting outcomes and referral rates for all New Zealand children. *NZMJ* 132: 9–19.
- Sand N, Silverstein M, Glascoe FP, et al. (2005) Pediatricians' reported practices regarding developmental screening: Do guidelines work? Do they help? *Pediatrics* 116: 174–179.
- SAS Institute Inc (2014) *SAS 7.1*. Cary, NC: SAS Institute Inc.
- Schluter PJ, Audas R, Kokaua J, et al. (2018) The efficacy of preschool developmental indicators as a screen for early primary school-based literacy interventions. *Child Development* 91: e59–e76.
- Schoenfeld D (1980) Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika* 67: 145–153.
- Shank L (2011) Mullen scales of early learning. In: Kreutzer JS, Deluca J and Caplan B (eds.) *Encyclopedia of Clinical Neuropsychology*. New York: Springer, pp. 1669–1671.
- Silva TB, Osório FL and Loureiro SR (2015) SDQ: Discriminative validity and diagnostic potential. *Frontiers in Psychology* 6.
- StataCorp (2017) *Stata Statistical Software: Release*. College Station, TX: StataCorp LLC.
- StataCorp (2019) *STATA survival analysis reference manual (15)*. <https://www.stata.com/manuals/st.pdf>
- Statistics New Zealand (2017a) How we keep IDI and LBD data safe. Available at: [http://archive.stats.govt.nz/browse\\_for\\_stats/snapshots-of-nz/integrated-data-infrastructure/keep-data-safe.aspx](http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/keep-data-safe.aspx) (accessed 31 October 2022).
- Statistics New Zealand (2017b) IDI Data Dictionary: B4 School Check Data. Available at: [http://archive.stats.govt.nz/browse\\_for\\_stats/snapshots-of-nz/integrated-data-infrastructure/idi-data/b4-school-check-data.aspx](http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/idi-data/b4-school-check-data.aspx) (accessed 2019).
- Statistics New Zealand (2019) Applying confidentiality rules to 2018 census data and summary of changes since 2013. Available at: <https://www.stats.govt.nz/methods/applying-confidentiality-rules-to-2018-census-data-and-summary-of-changes-since-2013> (accessed 10 January 2022).
- The Royal Children's Hospital Melbourne (2108) PEDS Brief administration and scoring guide. <https://www.rch.org.au/uploadedFiles/Main/Content/ccch/PEDS-Brief-Administration-and-Scoring-Guide.pdf>
- Thomas R, Sanders S, Doust J, et al. (2015) Prevalence of attention-deficit/hyperactivity disorder: A systematic review and meta-analysis. *Pediatrics* 135: e994–e1001.
- Wentz E, Björk A and Dahlgren J (2017) Neurodevelopmental disorders are highly over-represented in children with obesity: A cross-sectional study. *Obesity* 25: 178–184.
- Werner MJ, Walker LS and Greene JW (1994) Longitudinal evaluation of a screening measure for problem drinking among female college freshmen. *Archives of Pediatrics & Adolescent Medicine* 148: 1331–1337.
- WHO (2020) *Screening programmes: A short guide. Increase Effectiveness, Maximize Benefits and Minimize Harm*. Copenhagen: WHO Regional Office for Europe.
- WHO Anthro Department of Nutrition (2019) *Software for PC (Version 322, January 2011) and Macros [Internet]*. Geneva: WHO. Available at: <https://www.who.int/childgrowth/software/en/> (accessed 14 April 2020).
- Wikaire E, Curtis E, Cormack D, et al. (2016) Patterns of privilege: A total cohort analysis of admission and academic outcomes for Māori, Pacific and non-Māori non-Pacific health professional students. *BMC Medical Education* 16: 262.
- Williams AD, Clark TC and Lewycka S (2018) The Associations Between Cultural Identity and Mental Health Outcomes for Indigenous Māori Youth in New Zealand. *Frontiers in Public Health* 6: 316.
- Williams SJ (2013) *An exploration of nurses' experiences of delivering the before school check*. M. Phil Thesis, Massey University, Albany, New Zealand.
- Wilson JMG, Jungner G and World Health Organization (1968) *Principles and Practice of Screening for Disease*. Geneva: World Health Organization.
- Yusuf ZI, Dongarwar D, Yusuf RA, et al. (2020) Oral health problems among children with neurodevelopmental disorders in the United States. *International Journal of MCH and AIDS* 9: 157–160.
- Zablotsky B, Black LI, Maenner MJ, et al. (2019) Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics* 144: e20190811.
- Zhao J, Gibb S, Jackson R, et al. (2018) Constructing whole of population cohorts for health and social research using the New Zealand Integrated Data Infrastructure. *Australian and New Zealand Journal of Public Health* 42: 382–388.