National SCIENCE Challenges

A BETTER START

E Tipu e Rea

A Better Start

E Tipu e Rea

Brief Evidence Reviews for the Well Child Tamariki Ora Programme

Report submitted to MoH on 11 December 2019

Whakapūpūtia mai ō mānuka,
kia kore ai e whati

Cluster the branches of the manuka,
so they will not break

Foreword

The Ministry of Health is responsible for the development of policy advice on children's health and the future direction of the Well Child Tamariki Ora (WCTO) programme. The WCTO programme is the universal health service in New Zealand, which is responsible for protecting and improving the health and wellbeing of children from birth to 5 years of age. This is achieved through health and development screening and surveillance, whānau care and support, and health education.

The current programme is based on the evidence available at the time of the last programme update in 2007. Therefore, the Ministry of Health is reviewing the current WCTO Framework and associated Schedule (developed in 2002) to ensure that WCTO services meet the current needs of children and their whānau, and address the issues they face. The present review was initiated in 2019 and is the second review of the programme, as the first was carried out in 2006. In preparation for this review, the Ministry of Health has commissioned an evaluation of the recent literature on some of the new and emerging issues for preschool children, as well as possible ways to address them.

The purpose of this review includes ensuring that the programme is underpinned by the latest research and evidence. This is particularly pertinent to the current Schedule of Universal Contacts delivered, and one of the work-streams of the review is to consider the timing, content, and intensity of the Schedule, and associated additional contacts. This work stream will support the development of an integrated framework of universal wellbeing contacts for the pregnancy to 24 years of age life course.

The Ministry of Health require the brief evidence reviews (BERs) to synthesise relevant evidence about what works in key areas for children, including development, vision, hearing, emotional and mental health, and growth. The BERs adopted the He Awa Whiria – Braided Rivers approach and include consideration of what will work for Māori tamariki and whānau, and Pacific children and families within each domain. The BERs have helped to identify any knowledge gaps where further work and research may be needed, to inform further development of the WCTO programme.

The WCTO review is a key health contribution to the Government's Child and Youth Well-being Strategy. It forms part of the Ministry of Health's work programme to transform its approach to supporting maternal, child, and youth well-being.

The Ministry of Health have commissioned A Better Start: E Tipu E Rea National Science Challenge to undertake 11 health related BERs that will inform the WCTO review and decision making on the future core service schedule, and additional health and social services for children in New Zealand. The aim of the BERs is to ensure that decisions are grounded in, and informed by, up-to-date evidence. BERs are intended to synthesise available evidence and meet time constraints of health care decision makers. Internationally health technology agencies have embraced rapid reviews, with most agencies internationally offering these alongside standard reviews. These 11 BERs that we have conducted have been performed in a very short time which was a very challenging task.

A Better Start is a national research programme funded by the Ministry of Business Innovation and Employment (MBIE). The objective of A Better Start is to improve the potential for all young New Zealanders to lead a healthy and successful life. To achieve this, A Better Start is researching methods and tools to predict, prevent, and intervene so children have a healthy weight, are successful learners, and are emotionally and socially well-adjusted. A Better Start consists of more than 120 researchers across 8 institutions.

The BERs cover 11 domains critical to the WCTO programme, which are: neurodevelopment (#1); parent-child relationships (#2); social, emotional, and behavioural screening (#3); parental mental health problems during pregnancy and the postnatal period (#4); parental alcohol and drug use (#5); excessive weight gain and poor growth (#6); vision (#7); oral health (#8); adverse childhood experiences (#9); hearing (#10); and family violence (#11). The BERs have synthesised relevant evidence about what works in key areas for children across these domains, which were assessed with careful consideration of what will work for Māori tamariki and whānau and Pacific children and families. They have also identified knowledge gaps where further work and research may be needed to inform further development of the WCTO programme.

Within each domain, a series of 6–14 specific questions were drafted by the Ministry of Health, and subsequently refined with input from the large team of researchers assembled by A Better Start. A Better Start established discrete writing teams to undertake each BER. These teams largely consisted of a post-doctoral research fellow and specialty expert, often in consultation with other experts in the field. Subsequently, each BER was peer reviewed by at least two independent experts in the field, as well as two Māori and a Pacific senior researcher. In addition, senior clinical staff from the Ministry of Health have reviewed each BER. These were then revised to address all the feedback received, checked by the editors, and finalised for inclusion in this report.

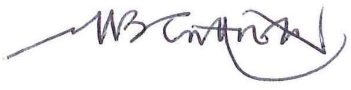
Whilst each of these domains are reviewed as discrete entities, there is considerably inter-relatedness between them. In particular, neurodevelopmental problems can be impacted by parent-child relationships, parental mental health, and pre- and postnatal drug exposure. Similarly, children who have problems with growth, vision, or oral health may also have neurodevelopmental disorders.

Most of the evidence available for these BERs comes from international studies with limited data from New Zealand, in particular there is limited information about Māori, Pacific, and disadvantaged families. These are the tamariki and whānau in whom the WCTO Programme services are more scarce, yet could potentially offer the greatest benefit.

The criteria for screening include the requirement for an effective and accessible intervention; the corollary is that screening should not be offered if there is no benefit to the individual being screened. The essential issue is therefore to identify those infants and preschool children and their whānau who would have better outcomes following intervention; this includes better outcomes for the whānau.

The current WCTO programme has had a greater emphasis on surveillance rather than screening. Many of the questions in the BERs address screening. A change in the WCTO programme that further extends into screening will require substantial upskilling of many WCTO providers, as well as redirection of resources. Importantly, Māori and Pacific iwi and community views must be considered before any new screening programmes are to be included.

It should be noted that a shift towards screening rather than surveillance may prevent health and behavioural problems. The economic benefits of prevention and early intervention are well documented, with early interventions showing that for every dollar spent there are substantial savings to health, social services, police, and special education resources.



Professor Wayne Cutfield
Director of A Better Start National Science Challenge
On behalf of the editors, authors and reviewers of the brief evidence reviews

National SCIENCE Challenges

A BETTER START

E Tipu e Rea

1 Neurodevelopmental screening and surveillance

Rajneeta Saraf BSc BBiomedSc MPH PhD[1]
Rosemary Marks MBChB FRACP[2]

[1] Department of Paediatrics, Child and Youth Health, University of Auckland, Auckland, New Zealand
[2] Starship Children's Hospital, Auckland District Health Board, Auckland, New Zealand

Suggested citation: Saraf R, Marks R. Neurodevelopmental screening and surveillance. In: Cutfield WS, Derraik JGB, Waetford C, Gillon GT, Taylor BJ [editors]. *Brief Evidence Reviews for the Well Child Tamariki Ora Programme.* A Better Start National Science Challenge. Auckland, New Zealand; 2019; p. 7-51

Table of Contents

List of Figures and Tables

Disclaimer

This brief evidence review was commissioned by A Better Start National Science Challenge (the Challenge) on behalf of the New Zealand Ministry of Health. It was prepared over a relatively short time based on the evidence available to the authors at the time of its preparation. The authors have made considerable efforts to perform a comprehensive and balanced evaluation of the existing evidence. However, this brief evidence review cannot be considered an exhaustive analysis of the existing peer-reviewed and grey literature on the topic, and it may not reflect the potentially conflicting views of all experts in the field. There could have been important omissions, and additional evidence might have also come to light since completion of this final draft. Thus, this brief evidence review should be considered with the appropriate caution. A previous version of this document was peer-reviewed by Māori and Pacific researchers and by independent experts in the field. Peer reviewers were anonymous, unless they have otherwise been identified by name. Please note that this brief evidence review does not represent the views of the Challenge or the Ministry of Health; rather, it reports the independent conclusions of the listed authors.

Conflicts of interest: The authors have no financial or non-financial conflicts of interest to declare that may be relevant to this work.

Abbreviations

AIMS	Alberta Infant Motor Skills
ASD	Autism spectrum disorder
ASQ	Ages and Stages Questionnaire
B4SC	B4 School Check
BDI	Battelle Developmental Inventory Screening Test
BOT	Bruininks-Oseretsky Test of Motor Proficiency
BSITD	Bayley Scales of Infant and Toddler Development
CAT/CLAMS	Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale
CP	Cerebral Palsy
DDST	Denver Developmental Screening Test
ELMS	Early Language Milestone Scale
FASD	Fetal alcohol spectrum disorder
GMs	General Movement Assessment
LDS	Language Development Survey
MABC	Movement Assessment Battery for Children
MAI	Movement Assessment of Infants
MCHAT	Modified Checklist for Autism in Toddlers – original and revised with follow-up versions
NDD	Neurodevelopmental disorder
NSMDA	Neurological Sensory Motor Development Assessment
NZ	New Zealand
PDMS	Peabody Developmental Motor Scales
PEDS	Parent Evaluation of Developmental Status
PLASTER	Paediatric Language Acquisition Screening Tool for Early Referral
PLC	Parent Language Checklist
SKOLD	Screening Kit of Language Development
SRST	Sentence Repetition Screening Test
TGMD	Test of Gross Motor Development
TIMP	Test of Infant Motor Performance

Executive Summary

- Limited evidence is available on the prevalence of neurodevelopmental disorders (NDD) in New Zealand children. Best estimates suggest a prevalence of between 3 and 10%; this is an underestimate for Māori and Pacific peoples. Having prospective cohort studies would be beneficial in providing robust national data on NDD prevalence and change over time. Identification of children with neurodevelopmental disorders is an important issue.
- Very limited information is available on the priorities for screening neurodevelopmental disorders. Expert consensus is that language development and hearing, FASD, ASD, Global Developmental Delay and Motor disorders (including cerebral palsy) are the top five neurodevelopmental screening priorities for New Zealand children under six years. Vision is also a priority but has been considered separately.
- One small study in Auckland has confirmed that Māori and Pacific children living in neighbourhoods of deprivation have a high incidence of neurodevelopmental problems[1].
- The current surveillance system using PEDS is not working for NZ Māori and Pacific peoples, and its use as a screening tool should be reviewed.
- Translation of the screening tools into commonly spoken languages in New Zealand e.g. Te Reo and Pacific Island languages and validation of these translated versions would prove to be beneficial for the culturally and linguistically diverse populations in New Zealand.
- There is a wide range of screening tools for use with children who may have neurodevelopmental problems. No one tool stands out as a comprehensive option for screening across the preschool age range for the wide range of neurodevelopmental problems. Tables of the sensitivities, specificities and utility of the various tools are provided.
- Families/whānau should receive information about screening so that they can make an informed decision about their child's participation.
- Screening processes need to be flexible to meet the needs of different populations.
- Timing of screening needs to be manageable for children/tamariki, families/whānau and screening providers. Therefore, information from all the domains covered by the review needs to be linked coherently. This needs to be collated and clear age points for screening identified.
- Potential harms of screening include inappropriate reassurance if screen is a false negative, and causing anxiety and stress if screen is a false positive. The failure of services to provide intervention in a timely way or through rationing of services is very stressful for families/whānau whose child has been identified as having a neurodevelopmental concern.
- Secondary screens may be appropriate when the primary screen has not provided a clear result. However, there needs to be rapid escalation to appropriate assessment and intervention when a significant deviation from normal development is identified.
- There is evidence that intervention is effective; choice of intervention for specific neurodevelopmental conditions is outside the scope of this review.

Abstract

Early development of motor and language skills is a useful indicator of a child's overall development and cognitive ability and is related to school success. Identification of young children at risk for developmental delay or related problems should lead to intervention services and family support, to ensure optimum opportunities for good outcomes. This evidence review was undertaken to evaluate the strengths and limits of primary and/or potential secondary screening and interventions for neurodevelopmental disorders, including motor dysfunction and speech & language delay in preschool-aged children, to determine the adverse effects of routine screening (if any) and what is known about screening in Māori and Pacific children. Studies reported wide ranges of sensitivity and specificity when compared with reference (sensitivity 22%-100%; specificity 55%-100%). The tools can be administered by a health professional, parent or a preschool teacher or a combination of any of these screeners. The shortest time to administer the screen was 5-10 minutes with some screens taking up to 60-90 minutes. It was found through this review that several aspects of screening have been inadequately studied to determine optimal methods, including which instrument to use, the age at which to screen, and which interval is most useful. PEDS Developmental Milestone (PEDS: DM), Ages and Stages Questionnaire (ASQ), and Brigance Early Childhood Screen may be used as secondary screening tests following a positive primary screen. No other evidence on secondary screening could be found. Interventional studies reported significantly improved motor and speech & language outcomes compared with control groups. However, the studies were small and long term effects are unknown. With the current surveillance system, Māori and Pacific peoples are underserved. Culturally appropriate approaches are needed to address this issue.

1.1 Introduction

Although there is no set definition for neurodevelopmental disorders (NDDs), it is often described as impairments in the functioning of the brain that affect a child's behaviour, memory or ability to learn. Examples of NDDs in children include attention-deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), learning disabilities, intellectual disability (previously known as mental retardation), conduct disorders, cerebral palsy, and impairments in vision and hearing. Genetic factors often contribute to these disorders. However, most NDDs are complex and have multiple contributors rather than one clear cause[2]. These disorders may likely result from a combination of genetic, biological, psychosocial, and environmental risk factors as well as behavioural risk factors. Environmental factors that may affect neurodevelopment include maternal use of alcohol, tobacco, and illicit and prescription drugs during pregnancy; prenatal or childhood exposure to some environmental contaminants; lower socioeconomic status; preterm birth; and low birthweight[3]. Children with NDDs often experience difficulties with language and speech, motor skills, behaviour, memory, learning, or other neurological functions[2]. While symptoms and behaviours of NDDs often change or evolve as a child grows older, some impairments are permanent. Identification of children at risk for developmental delay or related problems can lead to intervention services and family support at a young age when these are most likely to be effective[4]. Early intervention to address difficulties experienced by the child can reduce the risk or severity of certain types of neurodevelopmental disorders and improve developmental, emotional, academic and social outcomes.

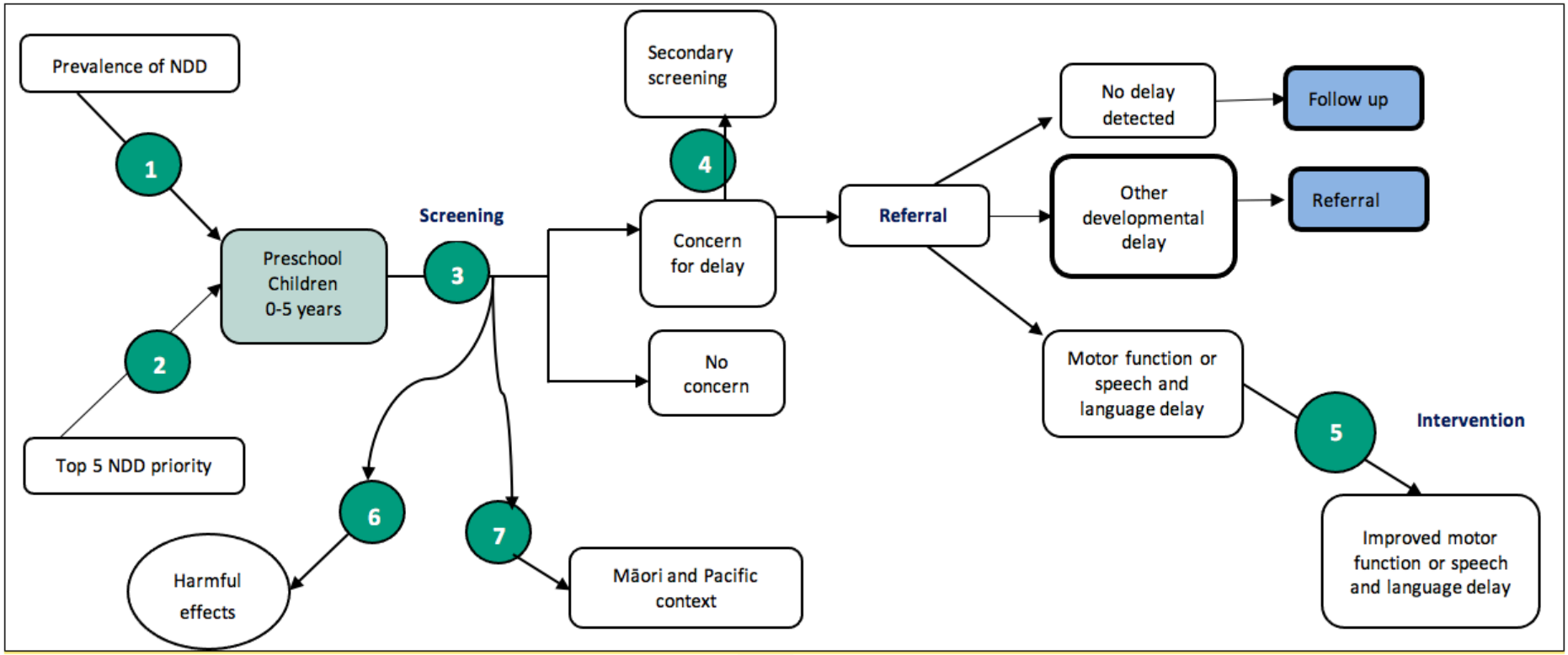
1.1.1 Key Questions

This review covers seven key questions (Figure 1.1). The key questions examine the evidence on the epidemiology of neurodevelopmental conditions including top five priority for NDDs in NZ children (key questions 1 and 2), about the effectiveness, accuracy and feasibility of screening children aged 5 years and younger for NDDs. (key question 3), secondary neurodevelopmental screening tests following a positive neurodevelopmental screen (key question 4), effectiveness of interventions for children identified with NDDs (key question 5), adverse or harmful effects of screening (key question 6), and screening from a Māori and Pacific perspective (key question 7).

1.1.2 Literature Search and Selection

Relevant studies were identified from multiple searches of MEDLINE (OVID), Embase and Cochrane library databases (1980 to August 2019). The following search terms were used: 'neurodevelopmental disorders', 'global developmental delay', 'disability', 'language delay', 'learning difficulty', 'intellectual difficulty', 'screening, surveillance', 'follow-up', 'referral', 'intervention', 'NZ children', 'NZ infants', 'NZ preschoolers', 'young children', 'under six years". Searches for each term were combined using Boolean operators. Articles were also obtained from recent systematic reviews[5-8], reference lists of pertinent studies, reviews, editorials, grey literature and by consulting experts. Some materials are not generally available and must be purchased, which limited the evidence review to published articles. All abstracts identified were reviewed and eligibility of full-text articles were determined based on several criteria. However; main criteria were availability of papers in English language, articles limited to screening / surveillance tools for gross motor, fine motor and language skills, and provision of primary data relevant to the seven key questions. A total of 86 full-text articles from searches and an additional six non-duplicate articles from reference lists met eligibility criteria and were reviewed. Data were extracted from each study and entered into evidence tables. Raw data was reported, and no statistical analyses were performed due to the heterogeneity of the studies.

Figure 1.1. The seven key questions representing an outline of the evidence review.
It includes the epidemiology and top five neurodevelopmental disorders in NZ children aged under six years, screening tools, interventions, adverse effects of the screening process and screening in Māori and Pacific children.



1. Prevalence of Neuro-Developmental Disorders (NDDs) in New Zealand children.
2. The top five NDDs that need screening in NZ children in early childhood (0 to 5 years).
3. Tests available to conduct primary NDD screening/ surveillance (accuracy, administration of screening instrument, associated costs and optimal time for screening).
4. Secondary NDD screening tests (if any) recommended following a positive screen and prior to an assessment.
5. Effective interventions following early detection and whether these interventions lead to significant improvements later in childhood / adolescence.
6. Any adverse / harmful effects from screening for an NDD during childhood.
7. What is known from a Māori and Pacific perspective about NDD screening in early childhood?

1.2 Prevalence of neurodevelopmental disorders in NZ children

It is challenging to accurately report the prevalence of neurodevelopmental disorders in pre-school aged children in New Zealand due to limited number of published papers in this area. Key sources include: Ministry of Health[9-11] & Statistics NZ Disability reports[12], publications from the Pacific Islands Families Study[13-16], Dunedin Multi-disciplinary Child Development Study[17], three studies in Hawkes Bay region[18-20], one by Gray[21] from the Counties Manukau Region, two in a provincial North Island city sampling 15 primary schools[22,23], and from a recent Master's project in Tamaki area[1]. The data from the Growing Up in New Zealand Study would have been valuable but was unavailable at the time this review was undertaken.

The Dunedin Multi-disciplinary Child Development Study reported 7.6 % and 10.4% of children born in 1972-1973 with language delays at 3 and 5 years of age respectively[17]. Results from the NZ Health Survey showed that 10.2% of children aged 3 to 4 years had emotional and behavioral difficulties in 2014-2015[11]. Two studies drawing data from the same cohort exploring outcomes for the B4 School Check (B4SC) in Hawkes Bay, found 7% of children (13% of referrals), had developmental concerns[18,19]. Another study on preschool children using the B4SC data in the Counties Manukau Region found that 3.4% had been identified with developmental concerns[21].

Data from the Pacific Islands Families Longitudinal Study depicts very high prevalence of developmental delays in 2-year olds (35%)[13], 16.8% internalising and 6.7% externalising behavioural problems in 2-year olds[14], and 26.9% of children having otitis media effusion at 2 years[15], with 2% of these children at increased risk of moderate to severe hearing loss at 11 years[16], The Welcome to School Study (WTS) in Tamaki, where 95% of children are Māori or Pacific, showed 22% of these children at 5 years, had developmental problems[1].

Across the five papers that reported the prevalence of developmental concerns or difficulties, the rate of variation was from 3.4 to 10.4%. There are limitations in using these data to estimate the burden of neurodevelopmental difficulties in NZ children due to the heterogeneity of the studies: prevalence is reported in different age groups, year of reporting is different and different assessment tools have been used. As demographic factors such as living in socio-economically deprived areas amongst others influence the rate of NDD, the prevalence of NDD is higher in Pacific children (6.7-35%) and the prevalence of 3.4-10.4% is an underestimate. It seems the prevalence rate in NZ preschool population may be somewhat similar to the USA (13.8%)[24]. More evidence is needed to provide more accurate prevalence data; this may become available from the Growing Up in New Zealand study.

1.2 Summary

Limited data is available on prevalence; based on the data that is available the prevalence of NDDs in New Zealand preschool aged children is between 3-10%. These values could be significantly underestimated as higher prevalence has been reported in Pacific children (6.7-35 %).

1.3 Priority for top five NDD screening in NZ children (0-5 years)

Only two data sources were found containing suitable data on NDD: NZ Disability 2013 survey[12] and Ministry of Health Report on the health of young people[25]. These sources were reviewed to determine the top five neurodevelopmental disorders that need screening in NZ children under six years of age. However, these sources contain limited and inconsistent data which is not specific to the 0-5 age group which is the focus of this review, but for a broader age range: 0-14 years. From these reports the priority of NDD screening is as follows:

1. Developmental delay including language delay, impaired social and cognitive skills, fine / gross motor skills
2. Psychology / psychiatric including behavioural, emotional and mental disorders
3. Physical impairments including cerebral palsy and other pervasive disorders
4. Intellectual disability
5. Hearing and Vision impairments

These priorities are in keeping with international literature. However in view of the lack of adequate data for the age range 0-5 years, the response to key question 2 was further developed in consultation with Developmental and Community Paediatricians (Dr Colette Muir and Dr Alison Leversha, personal communication 2019). Psychology / psychiatric including behavioural, emotional and mental disorders has been removed as this is covered in another Rapid Evidence Review (Domain 3). Vision is covered in Domain 7 of another Rapid Evidence Review. Our experts identified the conditions most commonly presenting to secondary care for further evaluations. The top five priorities for NDD screening in New Zealand children aged 0-5 years according to expert opinion are as follows:

1. Language development especially language deprivation, and including hearing screening
2. Fetal Alcohol Spectrum Disorder (FASD)
3. Autism Spectrum Disorder (ASD)
4. Global Developmental Delay; this term is preferred to Intellectual Disability as more appropriate to the preschool age group
5. Motor disorders including Cerebral Palsy (CP)

1.3 Summary

The top five NDDs that require primary screening in NZ children are language development and hearing, FASD, ASD, Global Developmental Delay and Motor disorders (CP).

1.4 Primary screening tools for NDDs, age of screening, cost, administration and accuracy

Even though we acknowledge that children with NDDs often experience an array of difficulties such as motor dysfunction, language & speech deprivation, problems with behaviour, memory, learning and other neurological functions, this brief evidence review mainly covers screening for motor skills (gross and fine) and speech and language delay. Key developmental abilities in the preschool years include vision, hearing, language, cognitive, social, emotional, and motor skills. Vision is considered in

Domain 7, and social, emotional, and behavioural development is discussed in Domain 3, therefore this review focusses on language (combined with hearing) and motor development and screening.

A total of 24 screening tools: 13 gross motor function and 11 speech & language delay screens have been identified through the literature search. An additional screen for Autism (M-CHAT) has been identified (Table 1.1):

Motor function screens

1. Alberta Infant Motor Skills (AIMS)[26,27]
2. Bruininks-Oseretsky Test of Motor Proficiency (BOT-2)[28]
3. Bayley Scales of Infant and Toddler Development (BSITD III)[29]
4. General Movement Assessment (GMs)[30-32]
5. Movement Assessment Battery for Children (MABC II)[33]
6. Movement Assessment of Infants (MAI)[34,35]
7. Parent Evaluation of Developmental Status (PEDS)[36]
8. Neurological Sensory Motor Development Assessment (NSMDA)[37,38]
9. Peabody Developmental Motor Scales (PDMS II)[39]
10. Test of Infant Motor Performance (TIMP)[40,41]
11. Test of Gross Motor Development (TGMD)[42]
12. Ages and Stages Questionnaire (ASQ)[36]
13. Infant Development Inventory (IDI)[43]

Social Communication, Speech & language screens

1. Battelle Developmental Inventory Screening Test (BDI II)
2. Battelle Developmental Inventory Screening Test (BDI II)Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale (CAT / CLAMS)[44]
3. Denver Developmental Screening Test – II (DDST II)[45]
4. Early Language Milestone Scale (ELMS)[46,47]
5. Fluharty Preschool Speech and Language Screening Test[48-50]
6. Language Development Survey (LDS)[51-53]
7. Levett-Muir Language Screening Test[54]
8. Parent Language Checklist (PLC)[55]
9. Pediatric Language Acquisition Screening Tool for Early Referral (PLASTER)[56]
10. Screening Kit of Language Development (SKOLD)[57]
11. Sentence Repetition Screening Test (SRST)[50,58]
12. Modified Checklist for Autism in Toddlers (M-CHAT) – original and revised with follow-up versions[59]

1.4.1 Reported accuracy of the identified screening instruments

Health professionals require standardised tools to identify, classify and diagnose developmental problems in children. Screening tools are either criterion referenced tests (the child passes if they achieve a specified criterion) or norm referenced tests (child's results are reported in relation to a specific population). The characteristics of the normed population should be considered as environmental and cultural differences have been found to affect development (motor).

These measurement features should be considered when selecting a developmental screener:

- Primary purpose: discrimination (normal vs abnormal), prediction (whether the child has a future risk of NDD or delay) or evaluation (monitor changes in development over time)
- Validity (content, construct and criteria- normally shown by factor analysis)
- Reliability (sensitivity, specificity, test-retest, inter, intra-reliability)
- Clinical utility (costs, time taken for administration, method of administration, screening age and whether training is required)

Validity and reliability characteristics are normally grouped as psychometric properties. Even though we acknowledge that all of the 4 characteristics stated above are important, the discussion on accuracy of the screening tool is mainly focussed on clinical utility and reliability (sensitivity and specificity).

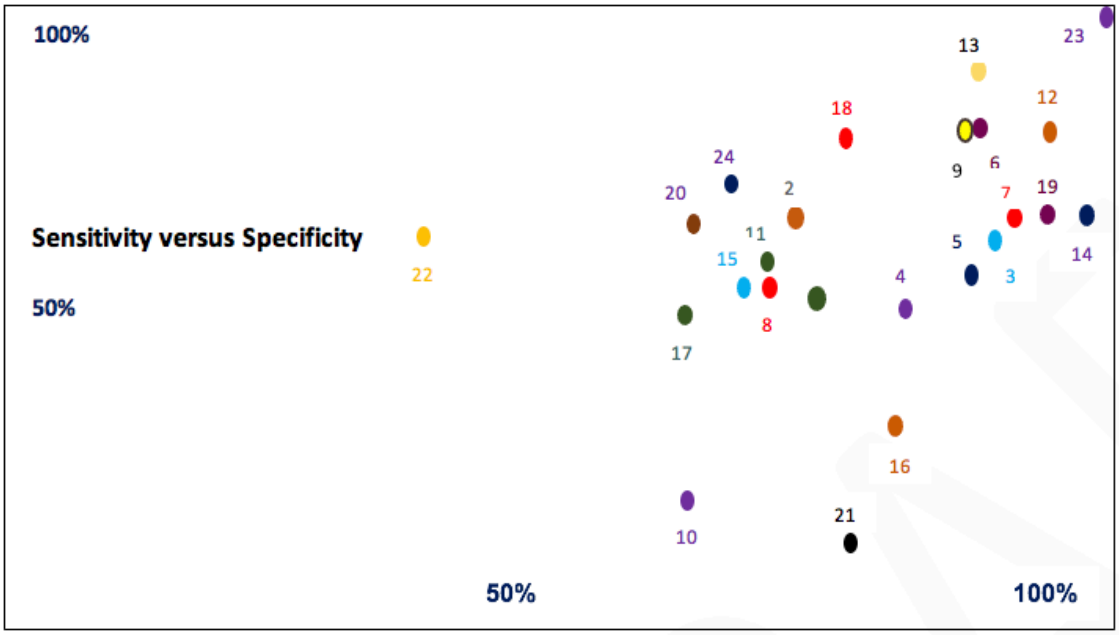
1.4.2 Psychometric properties

The sensitivity and specificity of most of the screening tools are good or excellent, except for Bruininks-Oseretsky Test of Motor Proficiency (BOT-2), Pediatric Language Acquisition Screening Tool for Early Referral (PLASTER), Denver Developmental Screening Test – II (DDST II) screens that have satisfactory sensitivity and specificity (Table 1.1 and Table 1.2 and Figure 1.2). Sensitivity ranged from 22% to 100% and specificity from 55% to 100%. Nine studies reported sensitivity and specificity of 80% or more using the General Movement Assessment (GMs)[29], Bayley Scales of Infant and Toddler Development (BSITD III)[30], Modified Checklist for Autism in Toddlers (M-CHAT)[41], Infant Developmental Inventory[43], Early Language Milestone Scale (ELMS)[46,47], Battelle Developmental Inventory Screening Test (BDI II), Language Development Survey (LDS)[51,52], the Clinical Linguistic and Auditory Milestone Scale (CLAMS)[53], and Levett-Muir Language Screening Test[54].

Studies utilising seven screening tools also provide evidence of the ability to discriminate between particular ages, which can be considered to support their content validity. The study of Neurological Sensory Motor Development Assessment (NSMDA) reported higher sensitivity / specificity at 8 months (83%/84%) compared to 1 (69%/73%), 4 (80%/57%), and 12 months (59%/94%)[37,38]. Similarly NSMDA[37,38], Test of Infant Motor Performance (TIMP)[40,41] and Movement Assessment of Infants (MAI)[34,35] screens have shown to have greater accuracy at 8 months, 9 months and 8 months respectively (Table 1.1).The study of the Clinical Linguistic and Auditory Milestone Scale reported higher sensitivity/specificity at age 14 to 24 months (83%/93%) than 25 to 36 months (68%/89%) for receptive function, but lower sensitivity/specificity at age 14 to 24 months (50%/91%) than 25 to 36 months (88%/98%) for expressive function[44]. A study testing expressive vocabulary using the Language Development Survey indicated higher sensitivity/specificity at age 2 years (83%/97%) than at age 3 years (67%/93%)[53]. The study of the Screening Kit of Language Development reported comparable sensitivity/specificity at ages 30 to 36 months (100%/98%), 37 to 42 months (100%/91%), and 43 to 48 months (100%/93%)[57].

For the motor function screens, studies have also reported the content validity and structural validity of BSITD III[29], BOT-2[28], Movement Assessment Battery for Children (MABC II)[33], Peabody Developmental Motor Scales (PDMS II)[39], Test of Gross Motor Development (TGMD)[42] to range from good to excellent, which indicated that these screens were actually measuring what they were supposed to[6,8].

Figure 1.2. The sensitivity and specificity of the screening tools.



1= AIMS: (73, 81.7), 2=GMs (83.3, 80), 3=NSMDA (58.8-82.4, 56.9-93.3), 4=PEDS (73-96, 73-86), 5=MABC II (79, 93), 6=TIMP (33-91.7, 75.7-94), 7=BSITD II (83, 94), 8= MAI (73.5-96, 62.7- 78.2), 9= PDMS (36.1-91.7, 52.3- 93.8), 10=BOT-2 (42.5-55.1, 65.7-72.6), 11= ASQ (82, 78), 12= MCHAT (91, 96), 13= ELMS (97, 93), 14= BDI-II (83, 100), 15= DDST II (73, 76), 16= PLASTER (53, 86), 17= SKOLD (73, 66), 18= LDS (91, 87), 19= CLAM (83, 97), 20= LDS (80, 67), 21= DDST (22, 88), 22=PLC (87, 47), 23=Levette- Muir Language Screening Test (100,100). 24=IDI (85, 77)

Table 1.1. The accuracy (sensitivity and specificity) of motor function screening tools.

Instrument	Age tested	Sensitivity	Specificity (%)	Outcome
AIMS[26,27]	0 to 18 mo	73.10%	81.7	Normal versus abnormal development Cerebral palsy
GMs[30-32]	0 to 4 mo	83.30%	80%	
NSMDA[37,38]	1 mo to 6 yr	68.8% (1 mo) 80.0% (4 mo) 82.4% (8 mo) 58.8% (12 mo)	72.6% (1 mo) 56.9% (4 mo) 83.7 % (8 mo) 93.3 % (12 mo)	Motor outcome- chance of walking
PEDS[60,61]	0 to 7–11 yr	73- 96%	73- 86%	
MABC-II[33]	3 to 16 yr	79%	93%	Motor impairment
TIMP[40,41]	32 wk to 4 mo	33% (1 mo) 50% (2 mo) 72% (3 mo) 62.5% (6 mo) 91.7% (9 mo)	94% (1 mo) 86% (2 mo) 91% (3 mo) 77.4% (6 mo) 75.7% (9 mo)	
BSITD-III[29]	1 mo to 3 yr	83%	94%	Motor impairment
MAI[34,35]	0 to 12 mo	73.5 (4 mo) 83.3 (4 mo) 96.0 (8 mo)	62.7 (4 mo) 64.5 (4 mo) 78.2 (8 mo)	Cerebral palsy
PDMS-II[39]	0 to 5 yr	36.1 (4 mo) 91.7 (8 mo)	93.8 (4 mo) 52.3 (8 mo)	Normal versus abnormal development
BOT-2[28]	4 to 21 yr	55.1 (4 yr) 42.5 (8 yr)	72.6 (4 yr) 65.7 (8 yr)	Motor Delay
MCHAT[59]*	16 to 30 mo	91%	96%	Autism
IDI[43]	0 to 6 yr	85%	77%	Motor delay

mo, month(s); wk, week(s); yr, year(s)
** M-CHAT is neither a motor nor a speech and language screener.*

Table 1.2. The accuracy (sensitivity and specificity) of language & speech delay screening tools.

Instrument	n	Reference Standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Ref
Parent Evaluation of Developmental Status	157	Clinical assessment	Expressive language, articulation	From outpatient clinic or private practice; 78% Caucasian; 54% male; 6-77 months	Clinic	Psychologist or Special education	72%	83%	[60,61]
Early Language Milestone Scale Clinical assessment	191	Clinical assessment	Expressive and receptive language	From private practices and pediatric outpatients of hospital; 80% Caucasian; 50% male; 0-36 months	GP clinic	Medical students	97%	93%	[47]
Early Language Milestone Scale	48	Receptive Expressive Emergent Language Scale, Bayley Scales of Infant Development	Expressive and receptive language	From low SES socioeconomic groups; 8-22 months	Pediatric clinic	Not reported	83%	100%	[46]
Denver Developmental Screening Test II (communication components)	89	Battery of measures	Fine motor, adaptive, personal social, gross motor, and language	From 5-day care centres; 52% male; 7-70 months	Day care centres	Psychologist	73%	76%	[45]
Pediatric Language Acquisition Screening Tool for Early Referral (PLASTER)	173	Early Language Milestone Scale	Expressive and receptive language	123 high risk infants; 50 normal controls; 3-36 months	*High risk:* neonatal developmental follow-up clinic *Control:* speech and hearing clinic	Speech & language pathologist	53%	86%	[56]
Bayley Infant Neurodevelopmental Screener	78	Bayley Scales of Infant Development II	Expressive and receptive language	Randomly selected from those presenting for routine neonatal high-risk follow-up; 54% male; 62% African American; 6-23 months	GP office	Developmental Paediatrician	73%	66%	[62]
Language Development Survey	306	Infant Mullen Scales of Early Learning	Expressive vocabulary	Toddlers turning 2- years old during the study in Wyoming; 52% male; 24-26 months	Home	Parent	91%	87%	[51]
Language Development Survey	64	Infant Mullen Scales of Early Learning	Expressive vocabulary	Children turning 2 years in a specific month in an area of Wyoming.	Home	Parent	83% (2 yr) 67% (3 yr)	97% (2 yr) 93% (3 yr)	[52]

Instrument	n	Reference Standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Ref
Language Development Survey	422	Bayley Scales of Infant Development, Stanford-Binet, Reynell Developmental Language Scales	Expressive vocabulary Delay 1: <30 words and no word combinations Delay 2: <30 words or no word combinations Delay 3: <50 words or no word combinations	Toddlers in four towns of Delaware County, PA turning 2-years old during the study	Home	Parent and research assistant	**Delay 1** Bayley 70%; Binet 52%; Reynell 67% **Delay 2:** Bayley 75%; Binet 56%; Reynell 89% **Delay 3** Bayley 80%; Binet 64%; Reynell 94%	**Delay 1** Bayley 99%; Binet 98%; Reynell 94% **Delay 2** Bayley 96%; Binet 95%; Reynell 77% **Delay 3** Bayley 94%; Binet 94%; Reynell 67%	[53]
Clinical Linguistic and Auditory Milestone Scale	99	Sequenced Inventory of Communication Development	Syntax, pragmatics	Infants turning 1 or 2 years old during study; 55% male; 0-36 months	Home or school for the deaf	Speech and language pathologist	**Receptive:** 14-24 months: 83% 25-36 months: 68% **Expressive:** 14-24 months: 50% 25-36 months: 88%	**Receptive:** 14-24 months: 93% 25-36 months: 89% **Expressive:** 14-24 months: 91% 25-36 months: 98%	[44]
Denver Developmental Screening Test II (communication components)	89	Battery of measures	Physical, self-help, social, academic, and communication	Children from five day care centres; 52% Male; 7-70 months	Day care centres	Psychologist	22%	86%	[45]
Parent Language Checklist	2,590	Clinical judgement	Expressive and receptive language	All children turning 36 months; 52% male; 41% urban	Home (mailed)	Parent	87%	47%	[55]
Structured Screening Test	376	Reynell Developmental Language Scales	Expressive and receptive language	Children from 2 low SES counties in London; Mean age 30 months	GP clinic	Health visitor	Severe: 66% Needs therapy:54%	Severe: 89% Needs therapy: 90%	[58]
Levett-Muir Language Screening Test	140	Reynell Developmental Language Scales, Goldman-Fristoe Test of Articulation, Language Assessment and Remediation Procedure	Receptive language, phonology, syntax	Private practice population; 34-40 months	GP Clinic	Medical practitioners	100%	100%	[54]
Fluharty Preschool Speech and Language Screening Test	279	Arizona Articulation Proficiency Scale Revised, Test of Language Development Primary	Expressive and receptive language, articulation	46% male; 74% Caucasian; 86% rural; 24-72 months	Preschool	Teacher	Speech & Language: 43% Speech: 74% Language: 38%	Speech & Language: 82% Speech: 96% Language: 85%	[50]

Instrument	n	Reference Standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Ref
Fluharty Preschool Speech and Language Screening Test	421	Test for Auditory Comprehension of Language Revised, Templin Darley Test of Articulation	Expressive and receptive language, articulation	52% male; 75% Caucasian; 24-72 months	Preschool	Teacher	Speech & Language: 31% Speech: 43% Language: 17%	Speech & Language: 93% Speech: 93% Language: 97%	[50]
Hackney Early Language Screening Test	1,205	Reynell Developmental Language Scales	Expressive language	Children attending routine developmental check-ups; mean age 30 months	Home	Health visitor	99%	69%	[63]
Fluharty Preschool Speech and Language Screening Test	90	Developmental Sentence Scoring	Expressive and receptive language, articulation	Children referred for speech and/or language assessment and intervention and controls; 24-72 months	Speech and hearing clinic in western Ontario	Clinician	10th percentile: 36% 25th percentile: 30%	10th percentile: 95% 25th percentile: 100%	[49]
Screening Kit of Language Development	602	Sequenced Inventory of Communication Development	Expressive and receptive language	From day care centres in Detroit; 30-48 months	Speech and language hearing clinic, day-care, GP clinic	Speech and language pathologists	30-36 months: 100% 37-42 months: 100% 43-48 months: 100%	30-36 months: 98% 37-42 months: 91% 43-48 months: 93%	[57]
Fluharty Preschool Speech and Language Screening Test	182	Sequenced Inventory of Communication Development	Expressive and receptive language, articulation	From day care programs; 36- 47 months	Clinic	Speech and language pathologists	60%	80%	[48]
Sentence Repetition Screening Test	76	Speech and Language Screening Questionnaire	Receptive and expressive language, articulation	Children registering for kindergarten; 48% male; 65% Caucasian; 54- 66 months	School	Non- specialists or school speech and language pathologists	Receptive and expressive: 62% Articulation: 57%	Receptive and expressive: 91% Articulation: 95%	[50]
Test for Examining Expressive Morphology	40	Kaufman Assessment Battery for Children, Structured Photographic Expression Language Test II	Expressive vocabulary, syntax	20 impaired and 20 unimpaired; 52% male; 73% Caucasian; 48- 67 months	School or clinic	Speech and language pathologists	90%	95%	[56]

1.4.3 Administration of the screening instruments (Table 1.3 and Table 1.5)

Most of the screening tools need to be administered by health professionals such as GPs, Paediatricians, Developmental Paediatricians, Nurses, Occupational or Speech and Language Therapists. Training is not needed to administer the screens that are conducted in clinical settings by health professionals; however, familiarity with the screen is required before administration. Some screening tools are completed by parents; these are the PEDS, ASQ, M-CHAT, IDI, LDS, and PLC and for the GMs screen, the child's video can be taken by the parent but all of these need to be scored by a health professional. For the parental-reported screens, it is important that parents are aware of developmental terms and milestones so that they are able to identify a developmental problem or concern; this is especially important for parents from culturally and linguistically diverse backgrounds. Two language and speech tools can be administered by preschool teachers (Fluharty Preschool Speech and Language Screening Test and SRST). Three screens can be performed in a school setting and teachers can assist the health professional with the screening of the child concerned (TGMD, MABC II, TIMP).

Table 1.3. Shows the method of administration of the screening tools.

Instrument	Administration of screen			Training
	Health professionals	Parents	Teachers	
AIMS	✔			x
GMs	✔	✔		✔
NSMDA	✔			x
TGMD	✔		✔	✔
PEDS		✔		x
MABC II	✔		✔	x
TIMP	✔	✔	✔	x
BSITD III	✔			✔
MAI	✔			✔
PDMS	✔			x
BOT-2	✔			x
ASQ		✔		x
MCHAT		✔		x
IDI		✔		x
BDI-II	✔			x
CAT / CLAMS	✔			x
DDST II	✔			x
ELMS	✔			x
Fluharty*			✔	x
LDS		✔		x
Levett-Muir**	✔			x
PLC		✔		x
PLASTER	✔			x
SKOLD	✔			x
SRST			✔	x

** Fluharty Preschool Speech and Language Screening Test.*
*** Levett-Muir Language Screening Test.*

1.4.4 Costs with each identified screening

Most of the assessment tools need to be purchased. The costs associated with purchasing these instruments range from $20 to $1650 and are provided in US dollars (Table 1.5). Prices could not be found for some tools. For three gross motor function screens (BOT-2, BSITD III, MABC II) and 1 language & speech screen (BDI-II), comprehensive kits need to be purchased containing examiner guides, manuals, scoring sheets, and activity equipment hence costs is high.

1.4.5 Optimal time or times to conduct screening test (Table 1.4 and Table 1.5)

Two screening instruments can only be used from birth to the first few months of life (GMs, TIMP), while two can be used from birth to the first year of life (MAI, CAT/CLAMS), and three from birth to the first few years of life (AIMS, PDMS, ELMS), and three from birth to beyond the preschool years (IDI, PEDS, BDI-II)

Table 1.4. Shows the reported age of screen administration.

Instrument	Birth to 6 months	6 months to 1 year	1 to 2 years	2 to 3 years	3-4 years	4 -5 years	5-6 years	>6 years
AIMS								
GMs								
NSMDA*								
TGMD								
PEDS								
MABC II								
TIMP								
BSITD III**								
MAI								
PDMS								
BOT-2								
ASQ								
M-CHAT								
IDI								
BDI-II								
CAT / CLAMS								
DDST II								
ELMS								
Fluharty								
LDS								
Levett-Muir								
PLC								
PLASTER								
SKOLD								
SRST								

** NSMDA recommended time of screening is from 1 month to 6 years (not from 6 months).*
**** BSITD III recommended time of screening is from 1 month to 3 years (not from 6 months).*

There are screening instruments that been used from the first few months of life to: (1) first few years of life (BSITD-III, M-CHAT, PLASTER), (2) 5 to 6 years of age (NSMDA), (3) beyond six years of age (DDST-II). Three tools have been reported to have been used past infancy (TGMD, MABC II, ASQ, Fluharty Preschool Speech and Language Screening Test, Levett-Muir Language Screening Test, PLC, SKOLD, SRST).

The administration time varied between different assessments with some studies noting that the older the child, the longer it takes for assessment[8]. PEDS, M-CHAT, ASQ, ELMS, Fluharty Preschool Speech and Language Screening Test, LDS, Levett-Muir Language Screening Test, PLC, PLASTER, SKOLD and SRST take the shortest time to administer (5-18 minutes). AIMS, GMs, and NMDSA, take 10-30 minutes, while TGMD-II and CAT/CLAMS are close with 15-20 minutes. TIMP and MABC-II take 20-40 minutes while IDI, BOT-2 and DDST-II take 20-30 minutes. The rest (BSITD III, MAI, PDMS II, BDI-II) take longer to administer (30-90 minutes).

1.4 Summary

The 25 assessment tools identified through the literature search are all appropriate for measuring motor development and speech & language delay in the preschool years. The most important step in identifying the best tool is to identify the purpose of the assessment and then choose a test that has been validated. One may wish to consider tools such as BSITD-III, PEDS, ASQ and DDST II that are appropriate to use for more than one function (motor function and language delay). Some tools such as GMs and BSITD-III require standardised training and may be costly, although this may improve the reliability and validity of screening. AIMS should be considered if an easy, accessible tool is needed that requires minimum handling and less time to administer.

Table 1.5. The clinical utility of the screening instruments identified through the literature search.

Tool	Short form	Age range	Time required (minutes)	Subscale measured	Method of Administration	Administrator	Costs
Alberta Infant Motor Skills[26,27]	AIMS	0-18 mo	20-30	Prone (21 items) Supine (9 items), Sitting (12 items) Standing (16 items).	Norm referenced. Therapist observes spontaneous activity in each of the subscales. Each Item is scored as least or most developmental mature, all items in between are marked as the "window period". Developmental maturity are scored as percentile scores.	Does not require specific training. Experienced therapists familiar with motor development and movement analysis are reliable testers. Non-therapists should receive training.	Scoring sheets are required. A pack of 50 sheets cost $48.95
Bruininks-Oseretsky Test of Motor Proficiency 2nd ed.[28]	BOT-2	4 yr to 21:11 yr	Complete: 45-60 Each composite: 10-15 Short: 15-20	Fine-motor precision, fine motor integration, manual dexterity, bilateral coordination, balance, running speed and agility, upper extremity coordination, strength.	Norm referenced. Clinician administered. Performance items including fine motor tasks, such as coping and tracing, and gross-motor tasks, such as sit-ups and running speed.	Preferably these tool should be administered by Paediatric health professionals, early childhood specialist. Formal training not required.	Comprehensive manual / kit: $1650. Test kit provides most equipment.
Bayley Scales of Infant and Toddler Development 3rd ed.[29]	BSITD III	1 mo to 3 yr	30-90	Assesses development in 5 areas: 1.Cognitive 2.Language 3.Motor 4.Social-emotional 5.Adaptive behaviour	The child is given tasks to measure cognitive skills, observed for receptive and expressive communication (language), assessed for motor skills, parental input is required for social-emotional and adaptive behavioural skills.	Preferably: Paediatric health professionals, early childhood specialist. Formal training not required. DVD, webinars and workshops available.	Comprehensive manual/ kit $1322. Test kit provides most equipment.
General Movement Assessment[30-32]	GMs	0 to 20 wk corrected	3-5 to video 20 for interpretation by trained professional	Gross movements, writhing movements, fidgety movements	General movements are assessed with the infant awake, lying on their back. The child should be calm and awake. The infant is videoed for 3-5 minutes and assessment is scored from the video.	Therapist, Allied health professionals can be trained to perform this assessment.	Comprehensive manual with DVD $80. Special video equipment needed.
Movement Assessment Battery for Children 2nd ed.[33]	MABC- II	3-6 yr 7-10 yr 11-16 yr	20-40	8 Tasks related to 3 specific areas: 1.Manual dexterity 2.Ball strikes 3.Balance (static and dynamic)	Assessment can take place at home, school or clinic. Movement is assessed in everyday situations. The examiner can assess groups of children in classroom situations, obtain parents or teachers views on child movement and measure the extent to which a child's attitudes and feelings about motor tasks are situation specific.	Can be performed by psychologists, speech therapists, physiotherapists, occupational therapists, mental health professionals, health practitioners, and education professionals. No additional specialised training is required.	Comprehensive manual / kit $1446. Test kit provides most equipment.

Tool	Short form	Age range	Time required (minutes)	Subscale measured	Method of Administration	Administrator	Costs
Movement Assessment of Infants[34,35]	MAI	0-12 mo	30-60	Four assessment domains: 1.Muscle tone 2.Primitive reflex 3.Automatic reactions 4.Volitional movement	Therapist observes and administers items in four assessment domains: 1.Muscle tone 2.Primitive reflex 3.Automatic reactions 4.Volitional movement	Physical therapist, occupational therapists, physicians, nurses, psychologists and others who have a good knowledge base in infant development. Special training to administer this exam is strongly recommended.	Cost not available.
Parent Evaluation of Developmental Status[36,60,61]	PEDS	0-7:11 yr	5-7	Testing items include questions on: 1.Development 2.Speech & Language 3. Learning & Cognition 4. Gross / fine motor skills 5.Social and emotional behaviour	10-item questionnaire that is completed by parents.	Health professionals. Training on how to administer PEDS Screen is offered.	Kit costs $66. Each kit has a scoring guide, 1 PEDS pad (x50) and scoring + interpretation form x50
Neurological Sensory Motor Development Assessment[37,38]	NSMDA	1 mo to 6 yr	10-30	Test items include: 1.Posture supine 2.Support on arms 3.Rolling 4.Prone Progression Creeping 5.Crawling hands and knees	The physiotherapist or clinician assess problems of posture, movement and coordination. An overall functional score is calculated in the grades in each of the 5 areas. Assessment forms available for ages: 1, 4, 8, 12, 18, 24, 36, 48, and 60 months.	Recommended for use in clinical setting therefore training in use of test is not essential but can be provided by accredited instructor.	Basic manual $20. Specific toys required but easily accessible.
Peabody Developmental Motor Scales 2nd ed.[39]	PDMS II	0-5 yr	30-60	Composed of 6 sub-tests: 1.Reflexes (reaction to stimulus) 2.Stationary (stand still) 3.Locomotion (crawl, hop, run, jump 4.Object manipulation (throw, catch) 5.Grasping (ability to use hands) 6.Visual-motor integration	The screen is a combination of task-related activities in each of the 6 subsets and recording of observations by the examiner of the child while doing the tasks.	Anyone can administer as long as they have knowledge on gross and fine motor functions and they can get training in how to use the screen.	$530 for the kit which has manual / guide to administer and score booklets.
Test of Infant Motor Performance[40,41]	TIMP	32 wk to 4 mo	20-40	Tests include: 1.Head control in supported sitting 2.Postural control in supine position 3.Righting reactions during tilting 4.Side-lying 5.Postural control in standing	Consists of 42 items in the 5 test areas. The examiner observes infant and then administers elicited items in standardised procedures.	Examiners can be teachers, health professionals (Occupational Therapists, Physiotherapists and doctors). No formal training is required.	Comprehensive manual / kit $60. Test kit provides most equipment.
Test of Gross Motor Development 2nd ed.[42]	TGMD II	3 yr to 10 yr	15-20	The tool is made up of 12 skills / tasks in 2 subsets: 1. Locomotor (run, hop, jump, slide etc.) 2. Object Control (catch, throw, kick etc.)	Standardised procedure. The examiner observes and scores the tasks.	TGMD-2 be administered by special physical educators, psychologists, occupational therapists, or physical therapists. Training is recommended.	Complete TGMD II Kit includes manual and 50x record forms $155.

Tool	Short form	Age range	Time required (minutes)	Subscale measured	Method of Administration	Administrator	Costs
Infant Developmental Inventory[43]	IDI	0 to 18 mo	20-30	1. Social 2. Self-Help 3. Gross Motor 4. Fine Motor 5. Language	The parent observes the child and scores the 5 areas. Recommended to start scoring from half the child's age.	Administered by caregivers.	$45 for the questionnaire and development chart.
Parent Evaluation of Developmental Status[36,60,61]	PEDS	0-7:11 yr	5-7	Testing items include questions on: 1.Development 2.Speech & Language 3. Learning & Cognition 4. Gross / fine motor skills 5.Social and emotional behaviour	10-item questionnaire that is completed by parents.	Health professionals. Training on how to administer PEDS Screen is offered.	Kit costs $66. Each kit has a scoring guide, 1 PEDS pad (x50) and scoring + interpretation form x50
Neurological Sensory Motor Development Assessment[37,38]	NSMDA	1 mo to 6 yr	10-30	Test items include: 1.Posture supine 2.Support on arms 3.Rolling 4.Prone Progression Creeping 5.Crawling hands and knees	The physiotherapist or clinician assess problems of posture, movement and coordination. An overall functional score is calculated in the grades in each of the 5 areas. Assessment forms available for ages: 1, 4, 8, 12, 18, 24, 36, 48, and 60 months.	Recommended for use in clinical setting therefore training in use of test is not essential but can be provided by accredited instructor.	Basic manual $20. Specific toys required but easily accessible.
Peabody Developmental Motor Scales 2nd ed.[39]	PDMS II	0-5 yr	30-60	Composed of 6 sub-tests: 1.Reflexes (reaction to stimulus) 2.Stationary (stand still) 3.Locomotion (crawl, hop, run, jump 4.Object manipulation (throw, catch) 5.Grasping (ability to use hands) 6.Visual-motor integration	The screen is a combination of task-related activities in each of the 6 subsets and recording of observations by the examiner of the child while doing the tasks.	Anyone can administer as long as they have knowledge on gross and fine motor functions and they can get training in how to use the screen.	$530 for the kit which has manual / guide to administer and score booklets.
Test of Infant Motor Performance[40,41]	TIMP	32 wk to 4 mo	20-40	Tests include: 1.Head control in supported sitting 2.Postural control in supine position 3.Righting reactions during tilting 4.Side-lying 5.Postural control in standing	Consists of 42 items in the 5 test areas. The examiner observes infant and then administers elicited items in standardised procedures.	Examiners can be teachers, health professionals (Occupational Therapists, Physiotherapists and doctors). No formal training is required.	Comprehensive manual / kit $60. Test kit provides most equipment.
Test of Gross Motor Development 2nd ed.[42]	TGMD II	3 to 10 yr	15-20	The tool is made up of 12 skills / tasks in 2 subsets: 1. Locomotor (run, hop, jump, slide etc.) 2. Object Control (catch, throw, kick etc.)	Standardised procedure. The examiner observes and scores the tasks.	TGMD-2 be administered by special physical educators, psychologists, occupational therapists, or physical therapists. Training is recommended.	Complete TGMD II Kit includes manual and 50x record forms $155.
Infant Developmental Inventory[43]	IDI	0 to 18 mo	20-30	1. Social 2. Self-Help 3. Gross Motor 4. Fine Motor 5. Language	The parent observes the child and scores the 5 areas. Recommended to start scoring from half the child's age.	Administered by caregivers.	$45 for the questionnaire and development chart.

Tool	Short form	Age range	Time required (minutes)	Subscale measured	Method of Administration	Administrator	Costs
Ages and Stages Questionnaire[36]	ASQ	40-60 mo	12-18	Contains 30 items and is available for assessment at 4, 5, 8, 12, 16, 18, 20, 24, 30, 36, 48 months. 30 items covering 4 areas: 1. Gross motor skills 2. Fine motor skills 3. Problem solving 4. Personal-social skills	Parent completed questionnaire as a general developmental screening tool.	Allied Health professionals. Training is provided through the Publisher.	$199 for the complete ASQ system (questionnaires and user guide)
Battelle Developmental Inventory Screening Test 2nd ed.	BDI II	0 to 7:11 yr	40-60	5 developmental domains assessed in any order: 1. Adaptive (ADP) 2. Personal- Social (P-S) 3. Communication (COM) 4. Motor (MOT) 5. Cognitive (COG)	Test administrators will use 3 different formats to obtain information about each child: (1) structured activities for direct assessment, (2) observation of child's natural environment such as home, day-care or school and (3) interviews with parents, caregivers and / or teachers.		The cost is approximately $1200 foe the initial kit & set manipulatives. Additional scoring sheets can be ordered.
Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale	CAT/ CLAMS	0-36	18-30	Includes psychometrics and speech and language milestones. CAT: 19 age sets with 12 instruments and 57 items for visual motor skills. CLAMS: 19 age sets with 3 instruments up to 24 months and 4 instruments after 24 months, includes 43 items for language skills	The test is focused on expressive, receptive, and visual language, primarily through parent report with occasional direct testing of the child.	Speech or language therapist.	
Denver Developmental Screening Test - II	DDST II	2 wk to 6 yr	20-30	Domains include: 1. Language (39 items) 2. Fine motor-adaptive (29 items) 3. Personal-social (25 items) 4. Gross motor (32 items)	Administered in a standardised manner with fine motor- adaptive activities delivered first followed by language, personal-social and gross motor activities.	Designed to be used in a clinical setting by a variety of professionals.	
Early Language Milestone Scale	ELMS	0-36	1-10	43 items covering 3 areas: 1. Auditory expressive 2. Auditory receptive 3. Visual (expressive and receptive)	Responses are obtained from a combination of parental/caregiver report, examiner observation, and direct testing. This assessment has three sections: auditory expressive, auditory receptive, and visual.	Developed for use in pediatric clinical setting as a brief screen for language abilities in <3 years. Administered by speech and language specialists.	Complete kit with manual and record forms (x100) $398
Fluharty Preschool Speech and Language Screening Test		3 to 6:11 yr	10	35 items separated into 3 sections (A, B, C) including identification of 15 common objects (phoneme), nonverbal responses to 10 sentences (syntax), and imitation of 10 one sentence picture descriptions. Assess identification, articulation, comprehension, and repetition	Activities involve articulation, repeating sentences, following directions, answering questions, describing action and sequencing events. Teacher questionnaire is also available.	Easy to administer. Examiners can be trained on how to score the items.	Complete kit $212. Each kit has 2 manuals, 2 picture books, 25x record forms and 12 blocks

Tool	Short form	Age range	Time required (minutes)	Subscale measured	Method of Administration	Administrator	Costs
Language Development Survey	LDS	18-35 mo	10	310 words arranged in 14 semantic categories. Parents indicate which words their child has spoken and describe word combinations of 2 or more words that their child has used.	Uses parents' reports of vocabulary and word combinations to identify language delays.	Can be completed independently at home by a parent.	
Levett-Muir Language Screening Test		34-40 mo	5-6	1. Receptive 2.Language, 3. Phonology, 4. Syntax	Test is divided into 6 sections: 1) Comprehension - child is asked to pick toys from group. 2) Vocabulary - child's ability to name the toys. 3) Comprehension - using pictures child is required to respond to questions. 4) Vocabulary - child's ability to name what's in the pictures. 5) Comprehension & representation - child's ability to answer "what" and "who" questions. 6) Overall - child is asked to explain the detailed composite picture.	Health professionals can administer this screen in a clinical setting.	
Parent Language Checklist	PLC	36 mo	5	12 questions for parents about their child's receptive and expressive language including one question assessing hearing problems	It can be completed independently at home by the parents.	Parents	
Pediatric Language Acquisition Screening Tool for Early Referral	PLASTER	3-36 mo	5-10	Communication development milestones by age with 7 individual areas. Each area contains 10 questions (5 relate to receptive language and 5 expressive language.		Speech and language pathologist.	
Screening Kit of Language Development	SKOLD	2.5 to 4 yr	10	Vocabulary comprehension, story completion, sentence completion, paired sentence repetition with pictures, individual sentence repetition with pictures, individual sentence repetition without pictures, auditory comprehension of commands.		Allied professionals or language and speech therapists.	
Sentence Repetition Screening Test	SRST	54-66 mo	10 or less	15 sentences repeated one at a time by the child after demonstration by the tester.	In a school setting (kindergarten)	Non- specialists or school speech and language pathologists.	
Modified Checklist for Autism for Toddlers	MCHAT	16-30 mo	5-10	Most widely used Autism Spectrum Disorder (ASD) Tool. Used to identify impairments in social interactions and communication and the presence of repetitive and restrictive behaviours. Some children may benefit from a more through developmental and Autism screening.	Initially parent administered and if a positive screen is obtained- follow up screening is performed with a health professional. Scored by health professionals	Parent and / or health professional.	Free

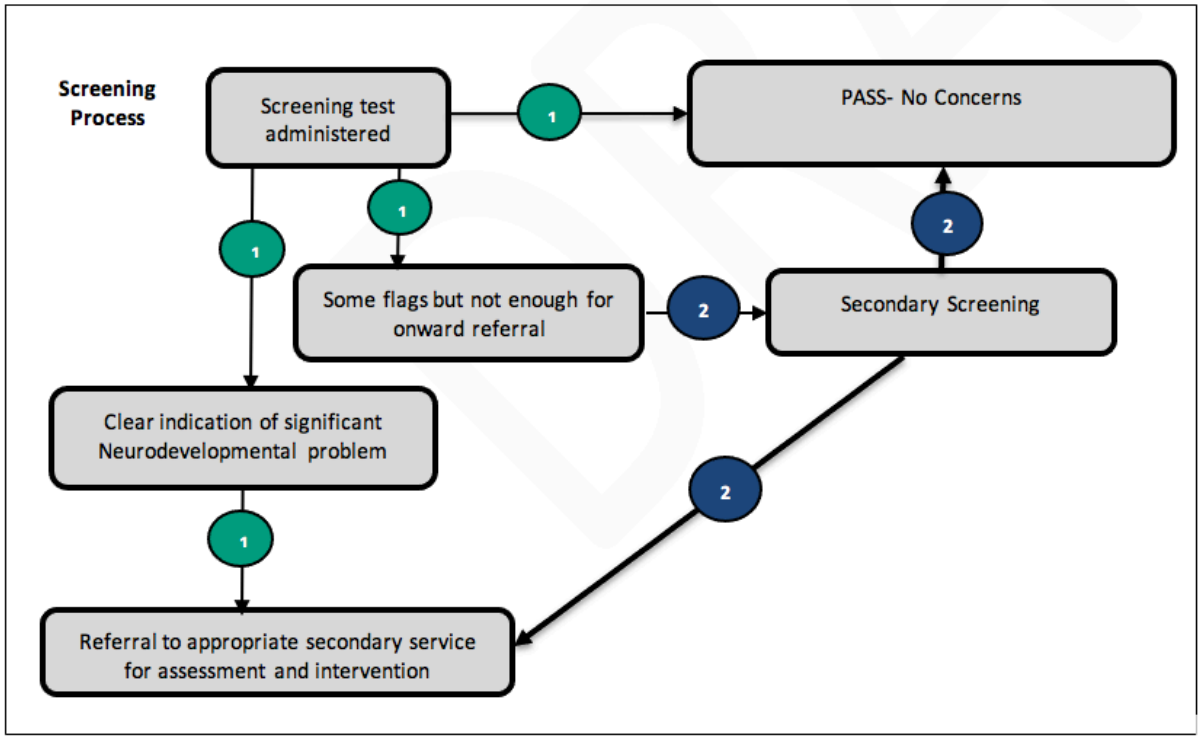
mo, month(s); wk, week(s); yr, year(s).

1.5 Secondary NDD screening tests following a positive screen

No international studies were found that addressed this question. On Royal Children's Hospital Melbourne website, there is information that if a concern is identified through PEDS as a primary screening instrument, a secondary screening tool may also be used as part of the assessment. The secondary screening instruments recommended are: PEDS: Developmental Milestones (PEDS: DM), the Brigance Early Childhood Screen and the Ages and Stages Questionnaire (ASQ).

According to Ministry of Health's recommendation, a secondary screen needs to be performed if a significant concern is highlighted in PEDS[1]. However not all B4SC staff have been trained to undertake this process. This was supported by two studies[21,64], who found that a second check for children identified with a potential issue was not offered. Wills (2010)[19] noted that in Hawke's Bay DHB, nurses were trained to conduct the Ages and Stages Questionnaire (ASQ) if predictive concerns were highlighted in the PEDS however; they did not report on whether this was done or on the outcome of the study[19].

Figure 1.3. The NDD screening pathway.



1.5 Summary

No evidence could be found on secondary NDD screening following a positive primary screen. Even though Ministry of Health recommends secondary screening following a positive PEDS screen, it is not clear whether this is done or how effective the process is. It should also be considered how a child who has been identified as having a potential developmental issue from the primary screen will benefit from a secondary screening when he/she should be directly referred to appropriate secondary services (Figure 1.2). A secondary screen should only be performed if some 'flags' are not enough for onward referral. This will prevent additional burden on the health care system.

1.6 Interventions leading to improved outcomes in early childhood

Twenty studies[65-82] were identified from the literature search as interventional studies; mainly randomised controlled trials. Studies were included if they used intention-to-treat analysis, method of randomization was reported, and there were more than 10 subjects in intervention or comparison groups. Limitations of studies, in general, include small numbers of participants (only 5 studies enrolled more than 50 subjects), lack of consideration of potential confounders, and disparate methods of assessment, intervention, and outcome measurement. As a result, conclusions about effectiveness are limited. Although children in the language and speech interventional studies ranged from 18 to 75 months of age, most studies included children aged 2 to 4 years old. Children in the motor function interventional studies were mostly older: 3 to 11 years (Table 1.6 and Table 1.7). Thus, the results do not allow for determination of optimal ages of intervention.

Studies evaluated the effects of individual or group therapy directed by clinicians and/or parents that focused on specific motor function (gross and fine) or speech & language domains. For motor function: these include locomotor, balance, object control and rhythm activities as well as activities on fine motor skills such as scissors cutting and shoelace tying. For speech and language domains: these include expressive and receptive language, articulation, phonology, lexical acquisition, and syntax. Several studies on speech & language delay, used established approaches to therapy, such as the HANEN principles[72,74]. Others used more theoretical approaches, such as focused stimulation[71,72], auditory discrimination[73], imitation or modelling procedures[77], auditory processing[63], and play narrative language[75,80]. Some interventions focused on specific words and sounds, used unconventional methods, or targeted a specific deficit.

Outcomes were measured by subjective reports from parents[71,72,75] and by scores on standardized instruments, such as the Reynell Expressive and Receptive Scales[71,76], the Preschool Language Scale[70], the MacArthur Communicative Development Inventories[75], and motor function scores obtained from MABC, PDMS II and NSMDA[65-69,83]. The most widely used outcome measure for language and speech improvement was mean length utterances, used by 3 studies[71,75,82] and object control and locomotor function used by 3 studies reporting motor function improvement[66,67,69].

A 12-month intervention (10-minute weekly sessions) in 18-42 months children as a treatment for receptive auditory comprehension led to significant improvement for the intervention group compared with control group, however, results did not differ between groups for several expressive and phonology outcomes[70]. Four studies evaluated speech and language interventions for children who were 2 to 3 years old[71,72,74,75]. Studies reported improvement on a variety of communication domains including clinician-directed treatment for expressive and receptive language[75], parent-directed therapy for expressive delay[71,72] and clinician-directed receptive auditory comprehension[70]. In 2 studies, there were no between group differences for clinician-directed expressive[70] or receptive language therapy[70], or parent-directed phonology treatment[74]. Five studies reported significant improvements for children 3 to 5 years old undergoing interventions compared with controls[70,76,77,79,80]. For motor function interventional groups, significant improvements were observed for balance, object control and locomotor function[65-69,83].

1.6 Summary

In general, studies of interventions were small and heterogeneous, may be subject to plateau effects, and reported short-term outcomes based on various instruments and measures. As a result, long-term outcomes are not known, interventions could not be compared directly, and generalization is questionable.

Table 1.6. Interventional studies to improve gross motor functions.

Motor skill assessment tool (s)	n	Age	Intervention Type	Intervention frequency and duration	Primary Outcome measures (s)	Results
Movement Assessment Battery for Children (MABC)[69]	76	5-8 yr	Kinder-kinetics-in-training (perceptual activities-locomotor, rhythm, balance and laterality-unilateral, bilateral and cross-lateral activities.	30 mins / 2 times per week for 8 weeks.	MABC-2: Manual dexterity, aiming and catching, balance	Balance increased in exposed group (p=0.05), whereas manual dexterity (p=0.797), aiming and catching (p=0252), showed no significant changes.
Peabody Development Motor Scales (PDMS II)[66]	149	54 mo	Skill based lesson plans were specifically designed to target stationery, locomotion, object manipulation, grasping and visual-motor integration skills for children in the experimental group. 16x lessons to target gross motor and 16x for fine motor.	16 weeks of 50 mins motor intervention (e.g. 25 min fine motor and 25 min gross motor)	Gross and fine motor skills	A repeated measure analysis of variance revealed a significant difference between the experimental and control group children on stationery (p<0.01) and visual-motor subsets (p<0.05) after the 16 weeks intervention.
Test of Gross Motor Development Assessment (TGMDA)[65]	27	3-5 yr	Parents tutored their children on academic readiness skills such letter, number, and colour recognition and on fine motor skills such as scissors cutting and shoelace tying.	Two 45-min lessons per week for 8 weeks delivered by the children's parents.	Gross and fine motor skills	The experimental group improved significantly in the object-control subscale score from pre-test to post-test (p<0.001), whereas the control group did not change.
Test of Gross Motor Development Assessment (TGMDA)[67]	59	4 yr	Skill intervention program consists of the following areas: 1. Hopping and galloping 2. Jumping 3. Ball bouncing 4. Striking 5. Kicking 6. Catching and throwing	24 instructional sessions (45 mins each) during a 12-week period.	Fine motor skills	Compared to the control group, the motor skill intervention group revealed significantly higher locomotor (p=0.000) and object control (p=0.000) scores following the intervention than prior to the intervention.
Test of Gross Motor Development Assessment (TGMDA)[68]	53	4-11 yr	The intervention group received the typical 'Successful Kinesthetic Instruction for Pre-schoolers" program and instructional motor skill program.	2 times a week for 9 weeks.	Locomotor and object control skills	The intervention group performed significantly better than the comparison group from pre to post-test for both locomotor (p<0.001) and object control skills (p<0.001).
Test of Gross Motor Development Assessment (TGMDA)[70]	40	4-5 yr	Each session consisted of: 1. A 2-3 min warm-up activity 2. 24 min of motor skill instruction for two object control skills 3. 2-3 min closure activity Two Mastery motivational climate (MMC) object control skills sessions were conducted each day.	30 mins session for 2 days per week for 9 weeks totalling 18 motor skill sessions.	Object control (OC) Perceived Physical Competence (PCC)	Both Object control skills and Perceived Physical Competence skills showed improvement after the 9 weeks intervention: PCC: p<0.001 OC: p<0.001

mo, months; yr, years

Table 1.7. Randomised controlled trials of interventions for speech & language delay.

Speech and language domains	n	Age (months)	Interventions	Speech and language outcome	Ref
Expressive and receptive language and phonology	159 in 2 groups	18-42	Clinician-directed individual intervention routinely offered by the therapist for 12 months vs. none	Improved auditory comprehension in intervention vs. control group; no differences for expressive language, phonology error rate, language development, or improvement on entry criterion	[75]
Expressive language	36 in 2 groups	27-39	Parent-directed individual therapy 60-75 minutes every other week for 6 months vs. none	Improved scores on several measures for intervention vs. control group	[73]
Expressive language	25 in 2 groups	23-33	Parent-directed individual focused stimulation intervention 150 minutes per week for 11 weeks vs. none	Larger vocabularies, use of more different words, more structurally complete utterances and multiword utterances in intervention group vs. control; no differences in several other measures	[74]
Expressive and receptive language	21 in 2 groups	21-30	Clinician-directed individual therapy 150 minutes per week for 12 weeks vs. none	Improved mean length of utterances, total number of words, lexical diversity, vocabulary size, and percentage of intelligible utterances in intervention group vs. control	[76]
Expressive language	25 in 3 groups	27-39	Clinician-directed individual therapy 60-75 minutes every other week for 6 months vs. parent-directed 60-75 minutes every other week for 6 months vs. none	Improved scores on all 5 measures for parent-directed group vs. control; improvement on 2 measures for clinician-directed group vs. control; improvement on 1 measure for parent vs. clinician group	[71]
Expressive language and lexical acquisition	10 in 2 groups	32-39	Clinician-directed individual therapy for 3 weeks vs. none	Improved multiword utterances from baseline in intervention group; no between group differences reported	[77]
Lexical acquisition and phonology	25 in 2 groups	23-33	Parent-directed individual therapy eight 150-minute sessions and 3 home sessions for 11 weeks vs. none	Improved level of vocalizations and inventory of consonants for intervention group vs. control; no differences in the number of vocalizations	[80]
Expressive and receptive language	39 in 2 groups	37-43	Clinician-directed interactive language therapy for 40 minutes weekly for 6 months (traditional group) vs. 40 minutes for 4 days per week for 3 weeks in two 3-month blocks (intensive group)	Improved expression score on Reynell scale for intensive group vs. weekly (or traditional) therapy group; no difference in comprehension scores, both improved	[78]
Expressive language	36 in 3 groups	47-83	3 clinician-directed approaches are compared for 5 months: mimicry, clinician modelling, 3rd person modelling for 5 months	Increased number of correct responses in modelling groups vs. mimicry group	[79]
Expressive and receptive language	30 in 3 groups	44-61	2 clinician-directed play groups with language impairments (treatment vs control) with normal peers for 20 minutes per week for 3 weeks	More words used, greater verbal productivity, more lexical diversity, and more use of linguistic markers by normal peer play group (not normal group, treatment group with language impairment) vs. control	[82]
Expressive and receptive language and phonology	159 in 2 groups	<42	Clinician-directed for 12 months vs none	Improved receptive language in intervention group vs. control; no differences between groups for 4 other measures	[70]

Speech and language domains	n	Age (months)	Interventions	Speech and language outcome	Ref
Phonology	26 in 2 groups	33-61	Clinician-directed individual therapy two 30-minute sessions per week for 4 months vs none	Higher scores on 3 of 4 measures for intervention vs. control group	[81]
Phonology	48 in 2 groups	50 (mean)	Clinician-directed individual therapy 30-40 minutes per week for 12 weeks; compares interventions for phonemes that differ (most knowledge/early developing group vs. least knowledge/latest developing group)	Improved scores on measures from baseline for both intervention groups; greater improvement for most knowledge/early developing phonemes group vs. comparison (least knowledge/latest developing) group	[84]
Phonology and syntax	26 in 3 groups	44-70	Clinician-directed sessions (individual and group) for 3 hours per week for 20 weeks vs. parent-directed sessions for 8 hours per week for weeks 1-12 (includes intensive parent training) then 4 hours per week for weeks 13-20 vs. none	Improved grammatical output (developmental sentence scores) for both intervention groups vs. control; no significant difference between groups for phonological output (percentage consonants correct)	[72]
Phonology	27 in 3 groups	42-66	Clinician-directed individual therapy 45 minutes per week for 6 weeks; compares 3 groups listening to different sets of words	45 minutes per week for 6 weeks; compares 3 groups listening to different sets of words Improved scores on measures for 2 intervention groups vs. third group	[73]
Syntax	28 in 3 groups	44-70	Clinician-directed vs. parent-directed vs. none for 5 months continuing from prior study	Improved some developmental sentence scores from baseline in both intervention groups vs. control; no between group comparisons reported, except that clinician-directed treatment groups had larger and more consistent gains than parent-directed treatment groups or control	[71]

1.7 Any adverse or harmful effects from screening for an NDD

No studies addressed this question. Potential adverse effects include false-positive and false-negative results. False-positive results can erroneously label children with "normal" development (speech, language and motor function), as impaired, potentially leading to anxiety for children and families and further testing and interventions. False-negative results would miss identifying children with impairment, potentially leading to progressive speech, language delay and motor function delay and other long-term effects including communication, social, and academic problems. In addition, once delay is identified, children may be unable to access services because they are past the specific age the interventions or services are targeted at.

Other adverse effects include the impact of time and cost of interventions on clinicians, parents, children, and siblings, loss of time for play and family activities, stigmatization, shame and labelling of the child and families with concerns or delays. Screening may also uncover a genetic disorder that has implications for other family members. There is also a risk that screening will identify more children than can receive intervention. This would be distressing for families/whānau and would create moral distress for clinicians and service providers.

1.7 Summary

One of the main adverse effects of NDD screening is the false-positive and false-negative test results which may cause anxiety and stress on the families and place addition burden on the health care system.

1.8 Screening from a Māori or Pacific perspective

In New Zealand, early detection of developmental and behavioural problems depends primarily on Well Child Tāmariki Ora providers (WCTO) with Parental Evaluation of Developmental Status (PEDS) used at 3 to 4, 5 to 7, 9 to 12, 15 to 18 months, 2-3 years and 4 years and Strength and Difficulties Questionnaire (SDQ) at 4 years. Both are being used as part of the "Before School Check (B4SC)" at 4 years. Recent data from four studies that were completed as part of the "Welcome to School (WTS)" study examining the health and development of children starting school in Tamaki, a multicultural community in Auckland where 95% of the children are Māori or Pacific, confirm the current developmental surveillance system using PEDS assessment tool is not working[1].

Study 1: Twenty out of the 93 children assessed, who had no concerns identified at the B4SC had concerns identified in the WTS study and, of which 13 were significant concerns[1].

Study 2: Reports that children starting school have low language skills which is a huge concern and parental reporting of language competence and language difficulties identified by PEDS, Early Childhood Education and Care (ECEC) and B4SC are not reliable[85]. The study identified the need for reliable language screening tools for NZ children, particularly those with Māori and/or Pacific heritage living in areas of deprivation.

Study 3: A third study exploring the nurses' perspectives on the B4SC indicates that the current utility of the B4SC is questionable and there is a need for better screening tools which are culturally appropriate, and which are delivered in a holistic manner[83].

Study 4: A fourth study evaluating whether the PEDS tool used in B4SC was achieving its purpose found that, of the 80% of children identified as having developmental concerns only 10.8% of these children were identified in the B4SC PEDS. The majority of those who were identified with the B4SC did not receive appropriate follow-up[86]. These findings suggest the PEDS which relies on parental concerns about development, may not be an effective tool for the NZ context; especially for Māori and Pacific peoples[86].

Cultural, linguistic or developmental literacy, or a combination of all three, are possible reasons for the inaccurate identification of children amongst Māori and Pacific peoples.

Cultural: Living in an area of high deprivation, where many children (more than 1 in 5 in WTS study[1]) demonstrate developmental delays, some parents may not have 'concerns' as comparisons are made with other children in their cultural groups who are developmentally similar. In addition, families/ whānau may be more accepting of difference and diversity than the predominantly European ethnic majority. They just accept that "Sione is Sione" (Dr Alison Leversha, personal communication 2019).

Linguistic: Many children with language delay from culturally and linguistically diverse backgrounds are not identified because their delay is attributed to bilingualism rather than impairment[87]. This applies to our Pacific peoples.

Developmental literacy: Difficulties have been reported with the administration of PEDS with families where English is a second language and / or literacy levels are low[1]. Very few parents reported concerns about their child's development at the B4SC and school entry, potentially signalling that among vulnerable communities such as Māori and Pacific communities, parents may not be aware of 'normal' development or have different understandings of development and are therefore less likely to recognise developmental delay[1,85,86]. Recognition of development and developmental delay is important as children living in a disadvantaged community during infancy are at increased risk of neurodevelopmental deficits (subtle problems in sensory motor and autonomic development that may be clinically unremarkable) but could interfere with child's adaptation and learning[87]. It has been noted that in the PEDS assessment there were hardly any children allocated to pathway D (parental difficulties understanding the questions) however in a population with high numbers of Pasifika families where English is a second language, the predicted numbers for pathway D could be higher[83].

PEDS as an assessment tool may not be culturally appropriate for the Māori and Pacific peoples. Despite being used in the WCTO schedule in NZ, PEDS has not been translated or validated for NZ populations. It would be beneficial if PEDS is translated into the commonly spoken languages in New Zealand e.g. Te Reo and Pacific Island languages.

1.8 Summary

With the current surveillance system, Māori and Pacific populations are underserved and there is an urgent need for culturally appropriate approaches. If the need for culturally appropriate approaches are not addressed, developmental concerns and delays in Pacific and Māori children will continue to be missed.

1.9 Conclusion

Studies are not available addressing the key question on recommended secondary screens following a positive neurodevelopment screen (key question 4), and adverse effects of screening (key question 6). Relevant studies are available regarding primary screening tools available for neurodevelopmental screening (key question 3), and effectiveness of early interventions on speech, language and motor function outcomes for children identified with delay (key questions 5) and screening in Māori and Pacific children (key question 7). Limited and inconsistent NZ studies were available to determine the prevalence and top five screening priority in New Zealand children under 6 (key questions 1 & 2).

Approximately 3-10% of New Zealand children under six years of age have neurodevelopmental disorders. However it is difficult to determine prevalence rates with accuracy as data is very limited and it's highly likely that this rate is under ascertained in Māori and Pacific peoples.

Language development and hearing, FASD, ASD, Global Developmental Delay and Motor disorders (CP) should be considered as the top five neurodevelopmental screening priority for New Zealand children under six years.

Fetal Alcohol Spectrum Disorder (FASD) is an umbrella term describing the range of physical, cognitive, behavioural and neurodevelopmental disabilities that can result from alcohol exposure during pregnancy. There is no NZ data on the prevalence of FASD, but international studies suggest that around 3% of births or around 1800 infants a year in NZ may be affected[88].

While maternal alcohol use should be picked up by screening for drug and alcohol use in pregnancy (covered in another Rapid Evidence Review-Domain 5), some women do not present to health providers until they are in labour. Also, we know that people often do not admit to alcohol use/amount of alcohol use. WCTO providers should be alert to the need to consider referral for secondary level assessment for FASD in a child who has poor growth and reduced head circumference; behavioural concerns; especially attention and retention of information; and/or developmental delay. So, we do not screen for FASD specifically but for developmental and behavioural issues with or without the context of poor growth.
A further question that needs to be considered by the Advisory Group is whether children whose mothers disclose alcohol use should be screened/monitored. Note that not all fetuses exposed to alcohol in utero develop FAS or FASD[88].

Vision is considered in another Rapid Evidence Review (Domain 7), so we have included hearing with language. There may be a case for consideration of hearing screening in a separate domain. The Newborn Hearing Screening should be offered to all newborn infants in New Zealand. Parents may decline screening, and some infants may miss screening for other reasons. Children with normal hearing at birth may develop hearing loss later as a result of middle ear disease or as the result of a congenital infection such as cytomegalovirus (CMV).

There is an overlap between neurodevelopmental concerns and behavioural concerns. Some children / tamariki with ASD may initially present with challenging behaviours. Conversely some children present primarily with developmental concerns which are the result of Adverse Childhood Experiences (ACEs). It is important that potential issues are picked up and routed to relevant services. Some re-routing between secondary services may be needed.

Although brief evaluations are available and have been used in a number of settings with administration by professional and nonprofessional individuals, including parents, the optimal method of screening for

motor skills and speech & language delay has not been established. Studies reported wide ranges of sensitivity and specificity when compared with reference standards (sensitivity 22% to 100%; specificity 55% to 100%). In these studies, the instruments providing the highest sensitivity and specificity included the General Movement Assessment (GMs), Bayley Scales of Infant and Toddler Development (BSITD III), Modified Checklist for Autism in Toddlers (M-CHAT), Early Language Milestone Scale (ELMS), Battelle Developmental Inventory Screening Test (BDI II), Language Development Survey (LDS), the Clinical Linguistic and Auditory Milestone Scale (CLAMS), and Levett-Muir Language Screening Test. Most of the evaluations, however, were not designed for screening purposes, the instruments measured different domains, and the study populations and settings were often outside primary care. No gold standard has been developed and tested for screening, reference standards varied across studies, few studies compared the performance of 2 or more screening techniques in 1 population, and comparisons of a single screening technique across different populations are lacking.

There is limited evidence on secondary screening so expert consensus input is needed. This is where the expertise of the primary screener becomes crucial; for example, midwives are expected to check the "red reflex" in newborn baby's eyes as a screen for congenital cataract. It is difficult to do this so there is a high rate of referral of false positives through to DHB ophthalmology services.

Randomised Controlled Trials of multiple types of interventions reported significantly improved motor function and speech & language outcomes compared with control groups. Improvement was demonstrated in several domains including object control, balance, locomotor function, articulation, phonology, expressive language, receptive language, lexical acquisition, and syntax among children in all age groups studied and across multiple therapeutic settings. However, studies were small, heterogeneous, may be subject to plateau effects, and reported short-term outcomes based on various instruments and measures. As a result, long term outcomes are not known, interventions could not be directly compared to determine optimal approaches, and generalizability is questionable.

There are many limitations of the literature relevant to screening for motor and speech & language delay in preschool-aged children including lack of studies specific to screening as well as difficulties inherent in this area of research. This evidence review is limited by use of only published studies of instruments and interventions. Data about performance characteristics of instruments, in particular, are not generally accessible and are often only available in manuals that must be purchased. Interventions vary widely and may not be generalizable. In addition, studies from countries with different health care systems, such as the U.K., and U.S may not translate well to NZ practice.

Although motor skill and speech & language development is multi-dimensional, the individual constructs that comprise it are often assessed separately. Numerous evaluation instruments and interventions that accommodate children across a wide range of developmental stages have been developed to identify and treat specific abnormalities of these functions. As a result, studies include many different instruments and interventions that are most often designed for purposes other than screening. Also, studies of interventions typically focus on 1 or a few interventions. In clinical practice, children are provided with individualized therapies consisting of multiple interventions. The effectiveness of these complex interventions may be difficult to evaluate. Adapting results of this heterogeneous literature to determine benefits and adverse effects of screening is problematic. Also, behavioural interventions are difficult to conduct in long-term randomized trials, and it is not possible to blind parents or clinicians. Randomizing children to therapy or control groups where clinical practice standards support therapy raises ethical concerns.

Identification of speech and language delay may be associated with benefits and adverse effects (mainly false positives / negatives) that would not be captured by studies of clinical or health outcomes. The

process of screening alerts physicians and caretakers to developmental milestones and focuses attention on the child's development, potentially leading to increased surveillance, feelings of caregiver support, and improved child self-esteem. Alternatively, caretakers and children may experience increased anxiety and stress during the screening and evaluation process. Therefore, it is important to consider whether counselling or appropriate and consistent information is offered to parents before screening. Detection of other conditions during the course of motor skill and speech & language evaluation, such as hearing loss, is an unmeasured benefit if appropriate interventions can improve the child's status.

1.10 Recommendations for further action

Policy and planning

- Translation of the screening tools into commonly spoken languages in New Zealand e.g. Te Reo and Pacific Island languages and validation of these translated versions would prove to be beneficial for the culturally and linguistically diverse populations in New Zealand.
- The current surveillance system using PEDS is not working for NZ Māori and Pacific peoples. A review of the current system is warranted to evaluate what is working and what is not, using this tool. Consideration should be given to the translation and validation of the PEDS tool in commonly spoken languages in New Zealand.
- Further policy work to determine the ages at which infants and children should be screened for NDDs should be coordinated with information from the Rapid Evidence Reviews for other domains. Screening instruments should be selected on the basis of the best ways of coordinating the varying screening processes.

Future research

- Future research should focus on determining optimal approaches of identifying preschool children with motor function and speech & language delay in primary care settings who would be appropriate candidates for further evaluations and possibly motor, speech & language interventions. These approaches should be integrated into routine developmental surveillance practices of clinicians caring for children.
- Studies that evaluate the effectiveness of validated brief screening instruments that include child and caretaker components could lead to a more standardized approach.
- Studies of specific motor, speech & language components of currently available broad developmental screening instruments, such as Ages and Stages Questionnaire, would be useful.
- Incorporation of risk factors and parent report in studies of screening approaches could provide information about their added value.
- Additional studies that compare screening instruments and methods in large primary care populations could lead to defining gold standards and acceptable referral criteria. Evaluating these criteria in different populations of children (e.g. Māori and Pacific) would minimize cultural and language biases.
- Additional work about the effectiveness of interventions, including motor, speech & language domain-specific results, may provide new insights.

- School-based efforts could be designed to complement strategies developed for young children improving long-term outcomes. Results of these studies may help determine optimal ages and intervals for screening. Functional long-term outcomes such as school performance, high school graduation rates, in-grade retention, special education placement/duration, and social adjustment need to be addressed more thoroughly.
- Cost-effectiveness evaluations of effective approaches that consider cost of treatment, the time that caregivers spend at treatment locations, the time they spend participating in the program on site or in the home, and long-term outcomes, among other factors, would be useful.

1.11 Graded evaluation of screening tools and interventions

We examined the strength and quality of evidence for neurodevelopmental outcomes to support the effectiveness of universal screening (Tables 1.8 and 1.9). Evidence found through a literature search was graded as “good”, “fair” or “poor” according to the definitions developed by the U.S. Preventive Services Task Force.[89]

For assessment of evidence for screening tools: a study was defined as “good” if a relevant available screening test was evaluated, a credible reference standard was used, reference standard was independently interpreted of the screening test; reliability of test was assessed and if the paper included a large sample size (more than 100) with a broad-spectrum patients. Evidence were treated as “fair” if relevant available screening test was evaluated, used reasonable although not best standard, reference standard was interpreted independent of screening test, had moderate sample size (50 to 100 subjects) and a “medium” spectrum of patients. “Poor” studies were those that had important limitations such as inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very small sample size of very narrow selected spectrum of patients.

For interventional studies- good studies were those where reliable and valid instruments were used, comparable groups were formed initially and maintained throughout the study, interventions were clear, important outcomes were considered and appropriate attention given to confounders in analysis. Others were categorised as poor or fair depending on the limitations.

For both screening tool and interventional studies, good studies were categorised as having high levels of certainty regarding net benefit; while fair studies as having moderate and “poor” studies as having low levels of net benefit.

Table 1.8. Graded evaluation of screening tools and associated recommendations for policy and practice.

Screening Tools	Grade	Estimated Net Benefit	Level of Certainty	Recommendation
Motor Function Screens				
Alberta Infant Motor Skills (AIMS)	C	Moderate	Moderate	This tool has been validated in large samples and cut-off established for abnormal motor development in 8-12 months old children. Therefore compares development with a norm-referenced group. It is an observational tool (takes 10-15 minutes to complete) so can be considered if there is a need for minimum handling.
Bruininks-Oseretsky Test of Motor Proficiency (BOT-2)	D	Moderate	Moderate	Evidence shows that BOT-2 is able to discriminate the motor development of infants as being normal or atypical. However the assessment is very lengthy-can take between 60-90 minutes and the scoring system is complicated. There is a shorter version of BOT-2 that takes 15-20 minutes but the correlation between the complete form and short version is not clear. Training is essential and all these need to be considered.
Bayley Scales of Infant and Toddler Development (BSITD III)	C	Moderate	Moderate	Evidence shows that BSITD III is the best practice tool for diagnosing developmental delay. BSITD III provides a comprehensive picture of the child's development (differentiate between receptive and expressive language, cognitive skills such as visual perceptual skills and play, fine motor manipulative skills, and gross motor skills). However an Australian study found that composite scores cannot be relied on for determining the degree of developmental delay (underestimates) and cultural issues alter the performance on individual items. Valuable tool if composite scores are revised and screen validated in common spoken NZ languages. However, takes 30-90 minutes to administer so may not be suitable for use with every child.
General Movement Assessment (GMs)	C	Moderate	Moderate	This tool can be offered to children who are at risk of neurodevelopmental disorders such as children born preterm, lack of oxygen, stroke or congenital heart disease. GMs is an observational tool and clinicians can be trained in the assessment technique. GMs tool may be valuable as evidence suggests that it can provide extra information on how a child's neurological system is developing.
Movement Assessment Battery for Children (MABC II)	C	Moderate	Moderate	Most commonly used tool to screen for motor function abnormalities. Can be used in three age groups: 3-6, 7-11 and 12-16. MABC-II can be used as an evaluative measure thus recommended for children in intervention programs (pre- and post- intervention) and if used for this purpose should be re-administered at a gap of at least 3 months from initial assessment. Evidence suggests that this tool may not be appropriate for certain ethnic groups for whom validation and translation may be required.
Movement Assessment of Infants (MAI)	C	Moderate	Moderate	Evidence shows that this tool provides the best information when administered to 4 month old infants. MAI can be offered to infants born at term who are at risk of neurodevelopmental delay. It can also help clinicians make decisions about intervention services.

Screening Tools	Grade	Estimated Net Benefit	Level of Certainty	Recommendation
Motor Function Screens				
Parent Evaluation of Developmental Status (PEDS)	B	High	High	PEDS is a simple, 10 item questionnaire completed by the parent and currently used as part of the WCTO programme in NZ. Evidence suggests that PEDS is a feasible developmental screening tool however three are concerns about the cultural appropriateness of PEDS; this needs further evaluation in a New Zealand context. It is highly recommended that PEDS be translated in common NZ languages and validated. Evidence also suggests that PEDS be used with secondary screening tool such as Parent Evaluation of Developmental Status: Developmental Milestones (PEDS: DM).
Neurological Sensory Motor Development Assessment (NSMDA)	B	Moderate	Moderate	Even though NSMDA can be used to assess motor development in children 1 month to 6 years of age, evidence shows the tool performs best at ages 8-12 months. Evidence shows that NSMDA measure is predictive-assessments done at early infancy can predict neurodevelopmental difficulties in preschool years (NSMDA measurements taken during infancy should be confirmed by another screen in the preschool years such as PEDsQL).
Peabody Developmental Motor Scales (PDMS II)	D	Moderate	Moderate	PDMS II screen is more complex and time consuming. Evidence shows that PDMS II is based on norm references. There is lack of agreement between development measures of PDMS II and BSITD III. Approximately half the children who showed appropriate total motor performance on the PDMS II were classified as delayed on the BSID II Motor Scale. Therefore PDMS should be used with caution or in combination with a second screen.
Test of Infant Motor Performance (TIMP)	B	Moderate	Moderate	There is fair evidence that TIMP provides a reliable and valid measurement that can be used for evaluation of motor function in term and preterm infants. Measurements are strongest in early infancy (aged 4 months or less). TIMP is highly reliable (highly sensitive and specific with the follow-up examination of BSID II) and has sufficient test-retest reliability. TIMP screen has the ability to discriminate among infants with differing risks for motor developmental delay. This screen can be recommended to all infants (risk or no risk).
Test of Gross Motor Development (TGMD)	B	Moderate	Moderate	There is fair evidence to say that TMGD II is reliable and appropriate assessment tool for assessing gross motor skill development of preschool children. The screen can be recommended for children with risk of neurodevelopment disorders as several validity studies have demonstrated TGMD's ability to differentiate children with cognitive impairments and autism spectrum disorder from typically developing children.

Screening Tools	Grade	Estimated Net Benefit	Level of Certainty	Recommendation
Motor Function Screens				
Ages and Stages Questionnaire (ASQ)	A	High	High	Evidence shows that ASQ are most cost effective, reliable way to screen children for developmental delays in the first 5.5 years of life. This parent completed screen has shown to correlate well with clinician's assessment. ASQ is been used worldwide and has been translated into many different languages. This will allow establishment of norm datasets from diverse ethnic groups. This screen can be recommended for use together with PEDS or SDQ.
Infant Development Inventory (IDI)	I	Moderate	Moderate	IDI is a brief questionnaire for use with children from birth to 18 months and takes approximately 10-15 minutes to complete. There is some evidence to indicate the accuracy of IDI- whether the tool correctly identifies children at risk for developmental problems (sensitivity) as well as accuracy with which the tool identifies the children not at risk. There is insufficient information to make any recommendations.

Screening Tools	Grade	Estimated Net Benefit	Level of Certainty	Recommendation
Social Communication, Speech & Language Screens				
Battelle Developmental Inventory Screening Test (BDI II)	C	Substantial	High	There is good evidence that this tool is effective in identifying children with a disability or developmental delay. However, the time taken to administer and cost need to be taken in to consideration in its use to screen the whole population.
Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale (CAT / CLAMS)	C	Moderate	Moderate	There is evidence that CLAMS could be used as a screening tool to detect children who have language delays quickly and easily. This tool can be considered for screening 1 to 3 year olds as it takes approximately 10 minutes to administer.
Denver Developmental Screening Test – II (DDST II)	D	Moderate	Moderate	There is fair evidence that this tool can be used to screen children in fine motor, adaptive, personal, social, gross motor and language domains and is able to detect children with or without problems. However this tool has high false positives.
Early Language Milestone Scale (ELMS)	C	Moderate	Moderate	There is fair to poor evidence that this tool is effectively able to identify children with expressive or receptive language difficulties and delays. This screen is recommended for children in the 2 to 3 years age group.
Fluharty Preschool Speech and Language Screening Test	D	Moderate	Moderate	There is evidence that Fluharty screen can be used to identify children with articulation impairments but the evidence suggests that Fluharty is too insensitive to be relied on for screening programs aimed at identifying preschool children with language disorders.
Language Development Survey (LDS)	C	Small	Moderate	There is good to fair evidence that the LDS screening tool has excellent sensitivity and specificity for identifying language delay at age 2 but somewhat lower levels for predicting developmental status one year later.

Levett-Muir Language Screening Test	I	Moderate	Moderate	Limited evidence presented was fair. The tool screens for receptive language, phonology and syntax. More evidence is needed to say whether this tool maybe suitable for NZ children.
Parent Language Checklist	A	Moderate	Moderate	There is good evidence that Parent Language Checklist may be used for prioritising children for referral to speech therapy services.
Paediatric Language Acquisition Screening Tool for Early Referral (PLASTER)	C	Moderate	Moderate	The evidence provided is fair. PLASTER is moderately successful in identifying children aged 3-60 months within normal limits for language development. Test-retest reliability was reported to be high. However sensitivity of PLASTER is poor.
Screening Kit of Language Development (SKOLD)	B	Moderate	Moderate	This tool has been validated in 2.5-4 year olds. There is fair evidence that this tool is able to identify a non-standard speaker from an impaired speaker. With NZ's diverse population, this tool may be important in identifying non-standard vs. impaired speakers once it has been translated and validated in common NZ ethnic populations.
Sentence Repetition Screening Test (SRST)	D	Moderate	Moderate	There is fair evidence that SRST tool is able to identify children with receptive, expressive and language articulation difficulties but the sensitivity of the tool has been reported as less than 70%. At this point this tool cannot be recommended for NZ preschool population.
Modified Checklist for Autism in Toddlers (M-CHAT)- original and revised with follow-up versions	A	Moderate	Moderate	From the evidence- M-CHAT revised version (M-CHAT-R) has shown to have greater utility than M-CHAT original. There is good evidence that M-CHAT-R detects Autism Spectrum Disorder (ASD) at a higher rate compared to M-CHAT and children who were diagnosed were 2 years younger than the national medium age of diagnosis. Implementation of M-CHAT-R as part of WCTO screening program can lower the age of ASD diagnosis by 2 years, increasing time for early intervention.

Grade: A, B, C, D, or I.
Estimated net benefit: substantial, moderate, small, nil or harmful, or insufficient (evidence).
Level of certainty: high, moderate, or low.
For more detailed explanation see Supplementary Information - Grade definitions and levels of certainty.

Table 1.9. Graded evaluation of interventions and associated recommendations for policy and practice.

Intervention	Grade	Estimated Net Benefit	Level of Certainty	Recommendation
Motor function interventions				
Motor Skill intervention program	I	Moderate	Moderate	Various types of movement-based interventions (balance / and or strength exercises, adapted play training, handball techniques, computerised games, a developmental physical education program, a therapeutic sensorimotor training programme, an intensive motor skills training programme, a physical therapy programme, and vestibular stimulation exercises) have shown to improve motor skills in children but the level of improvement differs from study to study. More evidence is needed through intervention studies to identify best motor function intervention, to examine sustainability of changes, and to examine the impact of intervention on other physical, health, social and emotional outcomes.
Parent assisted motor skills based intervention	B	Moderate	Moderate	Evidence shows that intervention outcomes could be enhanced if parents assist with the motor skills intervention program. Assistance could in the form of providing instructions during the program or home-based program delivery. Since the WCTO PEDS questionnaire is currently completed by parents. Having a parent-assisted motor skills intervention program could be considered for NZ children.
Teacher directed motor skills based intervention	I	Moderate	Moderate	Evidence suggests that interventions delivered by teaching staff maximises sustainability of the program, enhances participation and young children are more likely to be physically active when in school environment with peers. There is insufficient information as interventions of this sort places additional burden on teachers and are not usually encouraged. However teacher directed interventions could be undertaken in partnership with clinicians and researchers and may prove to be valuable.
Mastery Climate Motor Program	I	Moderate	Moderate	Insufficient information is available on whether mastery climate improves motor skills in children with developmental delays by increasing student engagement and addressing diverse learning needs of children. More evidence is needed.
Physical Activity or Language-enriched physical activity intervention.	B	Moderate	Moderate	Physical activity interventions such as Nintendo Wii Fit training, Martial arts training- Taekwondo, Trampoline and Table Tennis or language enriched physical activity intervention can be recommended to preschool children based on evidence available. However evidence suggests that physical activity motor skill programs should be underpinned by a sound theoretical framework.

Intervention	Grade	Estimated Net Benefit	Level of Certainty	Recommendation
Interventions for speech and language delay or disorder				
HANEN Approach	I	Moderate	Moderate	This could be parent or educator facilitated program to facilitate communication development in children. Targets language delays (It takes two to talk program), late talking (Target Word), Autistic Spectrum Disorder (More than words), and Asperger's (TalkAbility). There is moderate evidence that shows that benefits from HANEN intervention are similar to those from more traditional speech and language therapy. Insufficient evidence to make any recommendations for NZ children.
Imitation or Modelling	I	Moderate	Moderate	In this intervention program children were asked to mimic words vs. those who were taught grammatical rules. Fair-poor evidence suggests that different interventions work for different groups of children. Mimicry worked best in children with development impairment but teaching grammatical rules works better in children with typical development. More evidence is needed regarding this intervention.
Auditory language interventions	I	Moderate	Moderate	These are direct treatment approaches to influence children's ability to process speech and language such as speech-in-noise treatment, auditory recognition / discrimination, auditory system stimulation or modification of acoustic stimuli). There is lack of compelling evidence that auditory interventions would make significant contributions to auditory, language or academic outcomes of school aged children with auditory or speech and language impairment.
Verb focussed language intervention	C	Moderate	Moderate	Fair evidence shows that this intervention is effective in increasing the verb vocabulary of late talkers. It is not clear whether gains are sustainable over time.
Focussed Stimulation	B	Moderate	Moderate	A speech therapy where a child is asked to repeat a word or phrase multiple times in a conversation. Evidence shows that focussed stimulation improved child vocabularies and had a positive effect on language development. This intervention works well in children with expressive vocabulary delays or in late talkers. Vocabulary targets could be individually tailored for each toddler based on child's phonetic repertoire and parent report of vocabulary development.
Narrative Language Intervention	I	Moderate	Moderate	The intervention is provided in narrative language. Although the results presented in the papers were generally positive, each of the studies had limited number of participants, limited experimental controls and considerable variation in the methodology used. Insufficient evidence to make recommendations.

Grade: A, B, C, D, or I.
Estimated net benefit: substantial, moderate, small, nil or harmful, or insufficient (evidence).
Level of certainty: high, moderate, or low.
For more detailed explanation see Supplementary Information - Grade definitions and levels of certainty.

1.11 Summary

Clear recommendations can only be made in the context of future policy in relation to Well Child Tamariki Ora services. Screening tools can be divided into three broad groups: those that are completed by parental report, those that can be administered by people with minimal training and those which require specialist knowledge and training. The time taken to administer the various tools varies from 5 – 90 minutes. Tools vary in their sensitivity and specificity as well as with optimal age range for use, and all factors need to be considered in the context of timing of screening, workforce and access.

The evidence in relation to interventions is more challenging. As we discussed previously identification of a neurodevelopmental problem should lead to onward referral by the WCTO provider for verification (the secondary screen), in-depth assessment to ascertain the child's needs and establish the goals of intervention, and provision of an intervention programme to meet those needs. The population of children with neurodevelopmental problems is heterogenous with multiple aetiologies and trajectories. Therefore, comparisons are difficult.

One reasonably consistent group is children with cerebral palsy. Again, these children have multiple aetiologies for their impairment, and widely varying severity of impairment making comparisons difficult. Systematic reviews are available; however these become out of date rapidly because of development of new interventions.

The provision of interventions lies outside the current Well Child Tamariki Ora Framework, and policy formulation will need close collaboration with Child Development Services provided through Health and Early Intervention Services provided through Education.

References

1. Burge A. An examination of the development and wellbeing of children starting school in Tamaki, in Paediatrics. 2018, University of Auckland: Auckland.
2. Urkin J. Should we consider alternatives to universal well-child behaviour-development screening? Front Pediatr 2015;3:1-6.
3. Aarnoudse-Moens CSH, Weisglas-Kuperus N, van Goudoever JB, Oosterlaan J. Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. Pediatrics 2009;124:717-728.
4. Page F. Early detection of development and behaviour problems. Pediatr Rev 2000;21:272-280.
5. Sim F, Thompson L, Marryat L, Ramparsad N, Wilson P. Predictive validity of preschool screening tools for language and behavioural difficulties: A PRISAM systematic review. PLoS One 2019;14:1-31.
6. Spittle A, Doyle L, Boyd RN. A systematic review of the clinimetric properties of neuromotor assessments for preterm infants during the first year of life. Dev Med Child Neurol 2008;50:254-266.
7. Larson AL. Language Screening for infants and Toddlers: A literature Review of four commercially available tools. Communication Disorders 2016;38:3-12.
8. Griffiths A, Toovey R, Morgan P, Spittle A. Psychometric properties of gross motor assessment tools for children: a systematic review. BMJ 2018;8:1-14.
9. Simpson J, Duncanson M, Oben G, Adams J, Wicken A, Morris S, Gallagher S. The Health of Children and Young People with Chronic Conditions and Disabilities in New Zealand 2018, New Zealand Child and Youth Epidemiology Service, University of Otago: Dunedin.
10. Ministry of Health. Annual update of key results 2014/15: New Zealand Health Survey. 2015, Ministry of Health: Wellington.
11. Ministry of Health. Social, Emotional and Behavioural difficulties in New Zealand children- Technical Report: New Zealand Health Survey. 2018, Ministry of Health: Wellington.
12. Statistics New Zealand. Social and economic outcomes for disabled people: Findings from the 2013 Disability Survey. 2014, Statistics New Zealand: Wellington.
13. Paterson J, Iusitini L, Gao W. Child development assessment at 2 years of age: Data from the Pacific Islands Families Study. Pacific Health Dialog 2011;17:51-63.
14. Paterson J, Taylor S, Schluter P, Lusitini L. Pacific Islands Families (PIF) Study: Behavioural problems during childhood. J Child Fam Stud 2013;22:231-243.
15. Paterson J, Carter S, Wallace J, Ahmad Z, Garrette N, Silva PA. The Pacific Islands Study: Prevalence of chronic middle ear disease in 2-year old Pacific Children living in New Zealand. International Journal of Paediatric Otorhinolaryngology 2006;70:1771-1778.
16. Purdy SC, Taylor S, Schluter PJ, Tautolo E-S, Iusitini L, Ahmad Z, Sundborn G, Paterson J. Hearing and ear status of Pacific children aged 11 years living in New Zealand: the Pacific Islands families hearing study. Int J Audiol 2019;58:77-86.
17. Silva PA, McGee R, Williams SM. Developmental Language Delay from three to seven years and its significance for low intelligence and reading difficulties at age seven. Dev Med Child Neurol 1983;25:783-793.
18. Morris K, Wills R, Mara D, Stockdale A, Kirkpatrick S. Evaluation of the health outcomes for pre-school children associated with the Before School Check (B4SC) in the Hawke's Bay Region. 2010, Eastern Institute of Technology: Hawkes's Bay.
19. Wills R, Matthews K, Hedley C, Freer T, Morris H. Improving school readiness with the Before School Check: Early experience in Hawke's Bay N Z Med J 2010;123:45-58.
20. Hedley C. The B4 School Check behaviour measures: findings from the Hawke's Bay evaluation. Nurs Prax N Z 2012;28:13-23.
21. Gray S. Before School Check Audit. 2014, Counties Manukau District Health Board (CMDHB): Auckland.
22. Sargisson R, Powell C, Stanley P, de Candole R. Predicting motor skills from Strengths and Difficulties Questionnaire scores, language ability, and other features of New Zealand children entering primary school. Australian Educational and Developmental Psychologist 2014;31:32-46.
23. Sargisson R, Stanley P, de Candole R. Efficacy of quantitative screening assessments to identify new entrant children with potential difficulties. New Zealand Journal of Educational Studies 2013;48:66-81.
24. Woolfenden S, Eapen V, Williams K, Hayen A, Spencer N, Kemp L. A systematic review of the prevalence of parental concerns measured by the Parents' Evaluation of Developmental Status (PEDS) indicating developmental risk. BMC Pediatr 2014;14:231-236.

25. Craig E, McDonald G, Adams J, Reddington A, Oben G, Wicken A, NZ Child and Youth Epidemiology Service. The health of children and young people with chronic conditions and disabilities in NZ. 2011, Ministry of Health: Wellington.
26. Darrah J, Piper MC, Watt J. Assessment of gross motor skills of atrisk infants: predictive validity of the Alberta Infant Motor Scale. Dev Med Child Neurol 1998;40:495-491.
27. Piper MC, Darrah J, eds. Motor Assessment of the Developing Infant. 1994, Saunders, W B: Philadelphia.
28. Bruininks R, Bruininks B. Bruininks-Oseretsky Test of Motor Proficiency–2nd Edition (BOT-2): Manual, Pines C, Editor. 2005, AGS Publishing: Minnesota.
29. Bayley N. Bayley scales of infant development and toddler development: technical manual, PsychCorp T, Editor. 2006.
30. Cioni G, Prechtl H, Ferrari F, Paolicelli P, Einspieler C, Roversi MF. Which better predicts later outcome in full-term infants: quality of general movements or neurological examination? Early Hum Dev 1997;50:71-85.
31. Einspieler C, Prechtl HF, Bos AF, Ferrari F, Cioni G. Prechtl's Method on the Qualitative Assessment of General Movements in Preterm, Term and Young Infants, in Clinics in Developmental Medicine 2004, Mac Keith Press: London.
32. Ferrari F, Cioni G, Prechtl HF. Qualitative changes of general movements in preterm infants with brain lesions. Early Hum Dev Med Child Neurol 1990;23:193-231.
33. Henderson SE, Sugden DA, Barnett AL. Movement assessment battery for children-2, in Movement ABC-2: Examiner's manual. 2007, Pearson.
34. Chandler LS, Andrews MS, Swanson MW, Larson AH. Movement Assessment of Infants: A Manual. 1980, Rolling Bay: Washington.
35. Swanson MW, Bennett FC, Shy KK, Whitfield MF. Identification of neurodevelopmental abnormality at four and eight months by the movement assessment of infants. Dev Med Child Neurol 1992;34:321-327.
36. Limbos MM, Joyce DP. Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. J Dev Behav Pediatr 2011;32:1-13.
37. Burns YR, Ensbey RM, Norrie MA. The Neuro-Sensory Motor Development Assessment part 1: development and administration of the test. Aust J Physiother 1989;35:141-157.
38. Burns YR, Ensbey RM, Norrie MA. The Neuro Sensory Motor Developmental Assessment Part II: Predictive and concurrent validity. Aust J Physiother 1989;35:151-157.
39. Folio MR, Fewell RR. Peabody Developmental Scales. 2nd ed. 2000, Austin: Pro-ed.
40. Campbell SK. The Test of Infant Motor Performance, in User's Manual Version: Infant Motor Performance Scales. 2005, LLC: Chicago.
41. Campbell SK, Kolobe TH, Wright BD, Linacre JM. Validity of the Test of Infant Motor Performance for prediction of 6-, 9- and 12-month scores on the Alberta Infant Motor Scale. Dev Med Child Neurol 2002;44:263-272.
42. Ulrich DA. Test of gross motor development-2, Prod-Ed, Editor. 2000: Austin.
43. Creighton DE, Sauve RS. The Minnesota Infant Development Inventory in the developmental screening of high-risk infants at 8 months. Can J Behav Sci 1988;20.
44. Clark JG, Jorgensen SK, Blondeau R. Investigating the validity of the clinical linguistic auditory milestone scale. Int J Pediatr Otorhinolaryngol 1995;31:63-75.
45. Glascoe FP. Can clinical judgment detect children with speech-language problems? Pediatrics 1991;87:317-322.
46. Black MM, Gerson LF, Freeland CA, Nair P, Rubin JS, Hutcheson JJ. Language screening for infants prone to otitis media. J Pediatr Psychol 1988;13:423-433.
47. Coplan J, Gleason JR, Ryan R, Burke MG, Williams ML. Validation of an early language milestone scale in a high-risk population. Pediatrics 1982;70:677-683.
48. Allen DV, Bliss LS. Concurrent validity of two language screening tests. J Commun Disord 1987;20:305-317.
49. Blaxley L, Clinker L, Warr-Leeper GA. Two Language Screening Tests Compared with Developmental Sentence Scoring. Language, Speech, and Hearing Services in the Schools 1983;14:38-46.
50. Sturner RA, Funk SG, Green JA. Preschool speech and language screening: further validation of the sentence repetition screening test. J Dev Behav Pediatr 1996;17:405-413.
51. Klee T, Carson DK, Gavin WJ, Hall L, Kent A, Reece S. Concurrent and predictive validity of an early language screening program. J Speech Lang Hear Res 1998;41:627-641.
52. Klee T, Pearce K, Carson DK. Improving the positive predictive value of screening for developmental language disorder. J Speech Lang Hear Res 2000;43:821-833.
53. Rescorla L, Alley A. Validation of the language development survey (LDS): a parent report tool for identifying language delay in toddlers. J Speech Lang Hear Res 2001;44:434-445.

54. Levett L, Muir J. Which three year olds need speech therapy? Uses of the Levett-Muir language screening test. Health Visit 1983;56:454-456.
55. Burden V, Stott CM, Forge J, Goodyer I, The Cambridge Language and Speech Project (CLASP). Detection of language difficulties at 36 to 39 months. . Dev Med Child Neurol 1996;38:613-631.
56. Barnett LM, Minto C, Lander N, et al. Interrater reliability assessment using the Test of Gross Motor Development-2. J Sci Med Sport 2014;17:667-670.
57. Bliss LS, Allen DV. Screening Kit of Language Development: a preschool language screening instrument. J Commun Disord 1984;17:133-141.
58. Sturner RA, Heller JH, Funk SG, Layton TL. The Fluharty Preschool Speech and Language Screening Test: a population-based validation study using sample-independent decision rules. J Speech Hear Res 1993;36:738-745.
59. Robins DL, Casagrande K, Barton M, Chen CA, Dumont-Mathieu T, Fein D. Validation of the Modified Checklist for Autism in Toddlers, Revised With Follow-up (M-CHAT-R/F). Pediatrics 2014;133:37-45.
60. Glascoe F. Screening for developmenta; and behavioral problems. Ment Retard Dev Disabil Res Rev 2005;11:173-179.
61. Glascoe F, Marks K. Detecting children with developmental-behavioural probelms: The value of collaborating with parents. Psychological Test and Assessment Modelling 2011;53:258-279.
62. Macias MM, Saylor CF, Greer MK, Charles JM, Bell N, Katikaneni LD. Infant screening: the usefulness of the Bayley Infant Neurodevelopmental Screener and the Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale. J Dev Behav Pediatr 1998;19:155-161.
63. Law J. Early language screening in City and Hackney: The concurrent validity of a measure designed for use with 2-year-olds. Child Care Health Dev 1994;20:295-308.
64. Williams S. An exploration of nurses' experiences of delivering the Before School Check, in Nursing. 2013, Massey University.
65. Hamilton M, Goodway J, Haubenstricker S. Parent-assisted instruction in a motor skill program for at-risk preschool children. Adapted Physical Activity Quarterly 1999;16:415-426.
66. Hamilton M, Liu T. The effects of an intervention on the gross and fine motor skills of Hispanic pre-k children from low SES backgrounds. Early Childhood Education 2018;46:225-230.
67. Goodman J, Branta CF. Influence of a motor skill intervention of fundamental motor skill development of disadvantaged preschool children. Res Q Exerc Sport 2003;74:36-46.
68. Goodman J, Crowe H, Ward P. Effects of motor skill instruction on fundamental motor skill development. Adapted Physical Activity Quarterly 2003;20:298-314.
69. Lee D, Psotta R, Vagaja M. Motor Skills interventions in children with developmental coordination disorder: A review study. European Journal of Adapted Physical Activity 2016;9:20-29.
70. Robinson LE. Effect of a Mastery Climate Motor Program on Object Control and Perceived Physical Competence in Pre-schoolers. Res Q Exerc Sport 2011;82:355-359.
71. Fey ME, Cleave PL, Long SH. Two models of grammar facilitation in children with language impairments: phase 2. J Speech Lang Hear Res 1997;40:5-19.
72. Fey ME, Cleave PL, Ravida AI, Dejmal AE, Easton DL. Effects of grammar facilitation on phonological performance of children with speech and language impairments. J Speech Hear Res 1994;37:594-607.
73. Courtright JA, Courtright IC. Imitative modeling as a language intervention strategy: the effects of two mediating variables. J Speech Hear Res 1979;22:389-402.
74. Gibbard D. Parental-based intervention with pre-school language-delayed children. Eur J Disord Commun 1994;29:131-150.
75. Glogowska M, Roulstone S, Enderby P, Peters TJ. Randomised controlled trial of community based speech and language therapy in preschool children. BMJ 2000;321:923-926.
76. Robertson SB, Weismer SE. Effects of treatment on linguistic and social skills in toddlers with delayed language development. J Speech Lang Hear Res 1999;42:1234-1248.
77. Rvachew S. Speech perception training can facilitate sound production learning. J Speech Hear Res 1994;37:347-357.
78. Barratt J, Littlejohns P, Thompson J. Trial of intensive compared to weekly speech therapy in preschool children. Arch Dis Child 1992;671:106-108.
79. Cole KN, Dale PS. Direct language instruction and interactive language instruction with language delayed preschool children: a comparison study. J Speech Hear Res 1986;29:206-217.
80. Girolametto L, Pearce PS, Weitzman E. Interactive focused stimulation for toddlers with expressive vocabulary delays. J Speech Lang Hear Res 1997;40:338-348.

81. Almost D, Rosenbaum P. Effectiveness of speech intervention for phonological disorders: a randomized controlled trial. Dev Med Child Neurol 1998;40:319-325.
82. Robertson SB, Weismer SE. The influence of peer models on the play scripts of children with specific language impairment. J Speech Lang Hear Res 1997;40:49-61.
83. Williams S. An exploration of nurses' experience of delivering the Before School Check, in Nursing. 2013, Massey University [Masters Thesis].
84. Rvachew S, Nowak M. The effect of target-selection strategy on phonological learning. J Speech Lang Hear Res 2001;44:610-623.
85. Nelson M, Purdy S, Van Meygaarden A, Burge A, Leversha A. How do you know if a new entrant has significant language delay?, in Speech, Science and Psychology. 2018, University of Auckland [Masters Thesis].
86. Loo S, Burge A, van Meygaarden A, Nelson M, Purdy S, Ballard E, Kool B, Leversha A. New entrants in Tamaki: is the parent's evaluation of developmental status tool used in the before school check achiveing its purpose, in Paediatrics. 2017, University of Auckland [Masters Thesis]: Auckland.
87. Woolfenden S, Posada N, Krchnakova R, Crawford J, Gilbert J, Jursik B, V S, Perkins D, Kemp L. Equitable access to developmental surveillance and early intervention-understanding the barriers for children from culturally and linguistically diverse (CALD) backgrounds. Health Expect 2014;18:3286-3301.
88. Streissguth AP, Bookstein FL, Sampson PD, Barr HM. The enduring effects of prenatal alcohol exposure on child development: Birth through seven years, a partial least squares solution, in International Academy for Research in Learning Disabilities monograph series, No. 10, Press TUoM, Editor. 1993: US.

Supplementary Information - Grade definitions and levels of certainty

Table S1. Grade definitions for screening tools and interventions

Adapted with permission from the U.S. Preventive Services Task Force 2012.[i]

Grade	Definition	Recommendation for policy and practice
A	• The authors recommend this screening tool/intervention. • There is high certainty that the net benefit is substantial.	• This screening tool/intervention should be offered or provided.
B	• The authors recommend the screening tool/intervention. • There is high certainty that the net benefit is moderate, or there is moderate certainty that the net benefit is moderate to substantial.	• This screening tool/intervention should be offered or provided.
C	• The authors recommend selectively offering or providing this screening tool/intervention to patients based on professional judgment and patient preferences. • There is at least moderate certainty that the net benefit is small.	• This screening tool/intervention should be provided for selected patients depending on individual circumstances.
D	• The authors recommend against this screening tool/intervention. • There is moderate or high certainty that the screening tool/ intervention has no net benefit or that the harms outweigh the benefits.	• The authors discourage the use of this screening tool/intervention.
I	• The authors conclude that the current evidence is insufficient to assess the balance of benefits and harms of the screening tool/intervention. • Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.	• If the screening tool/intervention is offered, patients should understand the uncertainty about the balance of benefits and harms.

Table S2. Levels of certainty regarding net benefit

Adapted with permission from the U.S. Preventive Services Task Force 2012 [1].

Level Of Certainty	Description
High	• The available evidence usually includes consistent results from well-designed, well-conducted studies in representative populations. • These studies assess the effects of the preventive service on health outcomes. • This conclusion is therefore unlikely to be strongly affected by the results of future studies.
Moderate	• The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: – the number, size, or quality of individual studies; – inconsistency of findings across studies; – limited generalizability of findings to routine practice; – lack of coherence in the chain of evidence. • As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion(s).
Low	• The available evidence is insufficient to assess effects on health outcomes, because of: – the limited number and/or size of studies; – important flaws in study design and/or methods; – inconsistency of findings across individual studies; – gaps in the chain of evidence; – findings not generalizable to routine practice; – lack of information on important health outcomes. • More information may allow estimation of effects on health outcomes.

[i] https://www.uspreventiveservicestaskforce.org/Page/Name/grade-definitions