










SCOPING REVIEW

Clinical utility and psychometric properties of tools for early detection of developmental concerns and disability in young children: A scoping review

Andrea Burgess¹  | Carly Luke¹  | Michelle Jackman^{1,2}  | Jane Wotherspoon¹  | Koa Whittingham¹  | Katherine Benfer¹  | Sarah Goodman¹ | Rebecca Caesar³ | Tiffney Nesakumar¹ | Samudragupta Bora^{4,5} | David Honeyman⁷ | Danielle Copplin¹ | Sarah Reedman¹  | John Cairney⁶ | Natasha Reid⁸ | Leanne Sakzewski¹  | Roslyn N. Boyd¹ 

¹Queensland Cerebral Palsy and Rehabilitation Research Centre, Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia

²John Hunter Children's Hospital, Newcastle, NSW, Australia

³Women's and Children's Service, Sunshine Coast University Hospital, QLD, Australia

⁴Health Services Research Center, University Hospitals Research and Education Institute; Department of Pediatrics, University Hospitals Rainbow Babies and Children's Hospital, Case Western Reserve University School of Medicine, Cleveland, OH, USA

⁵Mater Research Institute, Faculty of Medicine & School of Psychology, Faculty of Health and Behavioural Sciences, The University of Queensland, QLD, Australia

⁶School of Human Movement, Faculty of Health and Behavioural Sciences, The University of Queensland, Brisbane, QLD, Australia

⁷Faculty of Medicine Library, The University of Queensland, Brisbane, QLD, Australia

⁸Child Health Research Centre, The University of Queensland, QLD, Australia

Correspondence

Andrea Burgess, The University of Queensland, Queensland Cerebral Palsy and Rehabilitation Research Centre (QCPRR), Centre for Children's Health Research, Level 6, 62 Graham Street, South Brisbane, QLD 4101, Australia.
Email: a.burgess@uq.edu.au

Funding information

University of Queensland, Grant/Award Number: Research Stimulus Grant (AB); Merchant Charitable Foundation via the

Abstract

Aim: To explore the clinical utility and psychometric properties of standardized tools for the early detection of developmental concerns or disability in young children.

Method: Systematic reviews and clinical practice guidelines containing psychometric data on tools appropriate for use with children from birth to 5 years 11 months were searched for in MEDLINE, CINAHL, Embase, and PsycINFO for the years 2000 to 2023, with no language restrictions.

Results: Eighty-six systematic reviews and six clinical practice guidelines guided identification of tools. A total of 246 tools were identified across domains of neurological,

Abbreviations: ABAS-3, Adaptive Behavior Assessment System, Third Edition; ADI-R, Revised Autism Diagnostic Interview; ADOS-2, Autism Diagnostic Observational Schedule, Second Edition; AIMS, Alberta Infant Motor Scale; ASD, autism spectrum disorder; ASQ-3, Ages & Stages Questionnaires, Third Edition; ASQ-TRAK, Ages & Stages Questionnaires - Talking About Raising Aboriginal Kids; BADDs, Brown Attention Deficit Disorder Symptom Assessment Scale; BASC, Behaviour Assessment System for Children; Bayley-III, Bayley Scales of Infant and Toddler Development, Third Edition; BDI-2, Battelle Developmental Inventory, Second Edition; BITSEA, Brief Infant Toddler Social and Emotional Assessment; BOT-2, Bruininks-Oseretsky Test of Motor Proficiency; BRIEF-P, Behavior Rating Index of Executive Function-Preschool; CARS, Childhood Autism Rating Scale; CBCL, Child Behaviour Checklist; CELF-5, Clinical Evaluation of Language Fundamentals, Fifth Edition; Conners EC, Conners Early Childhood Assessment; CSBS, Communication and Symbolic Behavior Scales; DAS-II, Differential Ability Scales, Second Edition; DCD, developmental coordination disorder; FASD, fetal alcohol spectrum disorder; GMA, Prechtl's General Movements Assessment; GMFM, Gross Motor Function Measure; HAI, Hand Assessment for Infants; HATS, Hearing and Talking Scale; HINE, Hammersmith Infant Neurological Examination; ITSEA, Infant Toddler Social and Emotional Assessment; MABC-2, Motor Assessment Battery for Children, Second Edition; M-CHAT-R/F, Modified Checklist for Autism in Toddlers, with Follow-up; MSEL, Mullen Scales of Early Learning; PDMS-2, Peabody Developmental Motor Scales, Second Edition; PEDI, Pediatric Evaluation of Disability Inventory; PEDI-CAT, Pediatric Evaluation of Disability Inventory - Computer-Adaptive Test; PEDS, Parents' Evaluation of Developmental Status; PLUM, Parents' Evaluation of Listening and Understanding Measure; PSEQ-H, Participation and Sensory Environment Questionnaire-Home; SACS-PR, Social Attention and Communication Surveillance Preschool; SACS-R, Revised Social Attention and Communication Surveillance; SDQ, Strengths and Difficulties Questionnaire; TGMD-2, Test of Gross Motor Development, Second Edition; VABS, Vineland Adaptive Behavior Scales; WPPSI-IV, Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition.

Plain language summary: <https://onlinelibrary.wiley.com/doi/10.1111/dmcn.16140>

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Developmental Medicine & Child Neurology* published by John Wiley & Sons Ltd on behalf of Mac Keith Press.

Children's Hospital Foundation, Grant/Award Number: Donation ID: 10415; National Health and Medical Research Council, Grant/Award Number: Early Career Fellowship (1145212 KB) and Research Fellowship (1105038 RB); Cerebral Palsy Alliance, Grant/Award Number: PHD02619 (CL); Department of Health and Aged Care, Australian Government

motor, cognition, communication/language, social-emotional, sensory processing, and/or specific diagnostic conditions of attention-deficit/hyperactivity disorder, autism spectrum disorder, cerebral palsy, developmental coordination disorder, and fetal alcohol spectrum disorder. After critical evaluation, 67 tools were included in the recommendations. Recommendations for screening and diagnostic assessment tools were based on best available evidence for predictive and discriminative validity, diagnostic accuracy, together with consideration of resource use and accessibility.

Interpretation: This comprehensive scoping review provides recommendations on the best tools for primary care, medical, allied health professionals, nursing, and other health workers to detect and identify developmental concerns or disability in young children using evidence-based tools.

Neurodevelopmental disorders are a group of conditions that manifest early in development, often before a child enters school. They can be characterized by developmental deficits, and they produce impairments of personal, social, academic, or occupational functioning (Diagnostic Criteria of Neurodevelopmental Disorders from the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision, DSM-5-TR).¹ Neurodevelopmental disorders considered in this review are those most commonly diagnosed, including intellectual developmental disorder (intellectual disability), global developmental delay, communication disorders, autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), cerebral palsy (CP), developmental coordination disorder (DCD), and fetal alcohol spectrum disorder (FASD).¹

Neurodevelopmental disorders frequently co-occur with one another.¹ Intellectual disability is the most commonly reported disability among children (e.g. 58% of Australian children who have a disability have an intellectual disability).² CP is the most common physical disability in childhood (1 in 700 children).^{3,4} The prevalence of autism is approximately 1 in 100 children.⁵ Definitions of neurodevelopmental disorders used in this review are those of the DSM-5-TR.¹

Early recognition of a developmental concern or disability is crucial for children and their families—helping families and carers to access appropriate supports and services that can facilitate their child's development and learning.⁶ Timely recognition of developmental concerns or disability in young children optimizes outcomes.⁶ Screening and assessment tools can support primary healthcare professionals to identify developmental concerns or disability at an early age, facilitating timely referral and access to intervention and supports. Screening tools are those that are used to identify children who are developing typically or who are at higher risk of developmental delays and are used to identify children who require more in-depth evaluation. Screening tools can be administered by a range of community and early childhood professionals. Diagnostic/assessment tools provide in-depth assessment of a child to determine the nature and extent of a developmental delay or to aid in diagnosis and are administered by appropriately trained practitioners.

What this paper adds

- Recommendations are provided for the most evidence-based and accessible screening and diagnostic assessment tools.
- The recommended multi-domain screening tool is the Ages & Stages Questionnaires.
- Recommended multi-domain developmental assessment tools are the Bayley Scales of Infant and Toddler Development, the Battelle Developmental Inventory, and the Mullen Scales of Early Learning.
- Recommended diagnostic assessment tools for cerebral palsy are Prechtl's Qualitative Assessment of General Movements and the Hammersmith Infant Neurological Examination.
- The recommended diagnostic screening tool for autism is the revised version of the Social Attention and Communication Surveillance.

Clinical practice guidelines highlight the importance of comprehensive multidisciplinary assessment for diagnosing neurodevelopmental conditions. A single screening or assessment tool cannot be used in isolation for diagnosis.

The aim of this scoping review was to provide primary healthcare professionals, namely general practitioners, paediatricians, nurses and midwives, allied health professionals (psychology, occupational therapy, physiotherapy, speech pathology), and community health workers, with information to inform the selection of the best evidence-based tools/assessments for the early detection of developmental concerns or disability in young children from birth to 5 years 11 months of age. Specifically, the review had the following aims: (1) to identify available assessments/tools for screening and diagnosis of children from birth to 5 years 11 months; (2) to provide recommendations on the most effective tools on the basis of available evidence for predictive and discriminative validity, diagnostic accuracy, and clinical utility.

METHOD

Searches

This scoping review of tools for the early detection of developmental delay or disability followed best practice guidelines as provided by the JBI Scoping Review network (<https://jbi.global/scoping-review-network/resources>).⁷ The process involved clearly defining the purpose and scope of the review; searching, identifying, and selecting studies; extraction/charting and categorizing data; and reporting and summarizing the results. The review was registered on Figshare (<https://doi.org/10.6084/m9.figshare.21708248.v1>) on 13th December 2022.⁸

Multiple approaches were used to search as broadly as possible and entailed the following: (1) a comprehensive search of relevant systematic reviews across four databases; (2) a comprehensive desktop search of grey literature sources (government and non-government, for example national and international clinical practice guidelines); and (3) specific searches on validation studies of tools identified in government guidelines and clinical practice guidelines that were not contained in the systematic reviews.

The comprehensive search of systematic reviews was undertaken across four electronic databases: MEDLINE ALL (Ovid), CINAHL (EBSCO host), Embase ([Embase.com](https://www.embase.com)), and PsycInfo (APA PsycNet). The search was limited to the years 2000 up to July 2023, and was not limited to the English language. The search strategies were developed by the first (AB) and second (CL) authors with the research librarian (DH). The search strategy used Medical Subject Headings (MeSH) terms and keywords for ‘children’, ‘disability’, ‘neurodevelopmental disorders’ or specific diagnoses; ‘assessments’ or ‘tools’; and ‘psychometrics’. To limit the search to systematic reviews, predefined methodological search filters were used. The MEDLINE ALL search was combined with the Search Strategy Used to Create the PubMed Systematic Reviews Filter (National Library of Medicine, 2018, adapted for MEDLINE Ovid), the CINAHL and Embase searches were combined with the Scottish Intercollegiate Guidelines Network systematic review search filters, and the methodology limit ‘systematic review’ was selected in PsycInfo. The search strategy is reported in Appendix S1. The search was initially performed on the 3rd August 2022 and updated on the 19th July 2023.

Systematic reviews were included if they met the following inclusion criteria: (1) contained tools/assessments that had published validity for children from birth to 5 years 11 months; (2) examined psychometric properties of tools (predictive, discriminative, or evaluative); (3) contained tools that assessed domains of motor (gross or fine), cognition, language/communication, behaviour, or social-emotional, and/or other neurodevelopmental outcomes; (4) contained tools used for diagnosing specific disabilities such as CP, autism, intellectual disability, DCD, and FASD; and (5) contained tools/assessments that could be parent-reported or patient-reported outcome measures

and/or standardized observational tools with normative data.

Systematic reviews were excluded if they met the following inclusion criteria: (1) were outside the age range (i.e. 6+ years); (2) did not examine tools that primarily assessed domains of motor (gross or fine), cognition, language/communication, sensory-motor, behaviour, or social-emotional or other neurodevelopmental outcomes; and (3) examined tools not available for use by primary healthcare professionals, allied health professionals, and community health practitioners (e.g. genetic testing, magnetic resonance imaging, tests exclusively performed in the neonatal setting).

Grey literature searches were performed by one author (AB) with guidance from the research librarian (DH). Google searches used a phrase (e.g. ‘diagnosis of autism in children’). The first four pages of search results were examined for relevant information. Relevant information was recorded in table form (Appendix S1). Targeted reference scanning and citation tracking of key papers were performed. Articles describing the tool were included if more than 90% of the study population were within the age range birth to 5 years 11 months.

Study selection of systematic reviews

Titles and abstracts of systematic reviews were independently screened by two reviewers (AB and CL) using Covidence. In cases of disagreement, a discussion was held and consensus reached. Full-text screening was performed by one reviewer (AB). The authors reviewed and confirmed tools and performed data extraction in teams by domain. As review teams performed data extraction, any full-text papers judged not to meet inclusion criteria were highlighted to the one reviewer (AB), discussed, and consensus reached about inclusion.

Data charting

Papers were categorized into tools that assessed domains of child development or tools related to specific diagnoses. Domains were neurological, cognition, communication and language, motor, sensory processing, social-emotional, and multi-domain tools and assessments. Diagnostic classifications were ASD, ADHD, FASD, CP, and DCD. Review teams comprised experts in each domain of child development or diagnoses to perform independent review and data charting.

Data charting was performed for tools/assessments from the included systematic reviews and clinical practice guidelines about psychometric properties and clinical utility (Appendix S2). Data were charted from papers in order of publication recency to avoid duplication of data from previous reviews. Systematic review summaries that used CONsensus-based Standards for the selection of health Measurement INstruments (COSMIN)⁹ were used when available. COSMIN guidelines⁹ were followed when reporting on measurement properties and quality of studies.

Identified tools/assessments were ranked within domains for recommendation on the basis of predictive and discriminative validity and clinical utility by the review team of experts.

RESULTS

Search results

Figure S1 shows the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) flowchart for the systematic review search. Eighty-six systematic reviews (Appendix S5), four national clinical guidelines,¹⁰⁻¹³ and two international clinical guidelines^{14,15} were included in this review. A total of 133 additional papers were sourced to provide further information on specific measures. A total of 246 tools were identified, of which 67 were appropriate for children aged from birth to 5 years 11 months, had published validity and reliability, and were available for use (Table 1; Appendix S2 and S3). Some tools assessed multiple domains of development, others were specific to a domain or diagnosis (e.g. autism), while yet other tools could be used across different domains and diagnoses (such as ADHD, behaviour, and social-emotional domain). Four tools were identified that had been specifically validated for Australian Aboriginal and Torres Strait Islander children.¹⁶⁻¹⁹

Multi-domain tools

Multi-domain assessment tools can be used to screen or assess a child across multiple developmental domains (e.g. cognition, motor skills, communication, etc.) or multiple functional/adaptive domains (e.g. self-care, mobility, socialization).

Multi-domain developmental screening tools

There was extensive psychometric evidence²⁰⁻²⁷ for the Ages & Stages Questionnaires, Third Edition (ASQ-3).^{28,29} The ASQ was recently recommended as the most appropriate multidimensional screening tool of child development.^{21,30} The ASQ-3 aims to identify children with concerns across cognition, motor, communication, and personal-social domains from 1 month to 5 years 6 months (1-66 months) of age. The ASQ is the most widely used developmental screening questionnaire in follow-up cohort studies.²⁵ A recent systematic review²³ found that, by using a threshold of more than 2 standard deviations (SD) below the mean, the ASQ had pooled sensitivity of 0.77 (95% confidence interval [CI] 0.64-0.86) and specificity of 0.81 (95% CI 0.75-0.86) to diagnose any developmental delay, and a sensitivity of 0.84 (95% CI 0.75-0.90) and specificity of 0.77 (95% CI 0.71-0.82) to diagnose severe developmental delay in children aged 12 to 60 months.²³

The ASQ is easy to use and can be administered through telehealth with parents/caregivers. The ASQ-3 is currently available in 11 languages. Cultural validation of the ASQ for Indigenous Australian children has led to the development of the Ages & Stages Questionnaires - Talking About Raising Aboriginal Kids (ASQ-TRAK).^{31,32}

There was evidence available for the Parents' Evaluation of Developmental Status (PEDS;³³ the revised version of the PEDS; PEDS-Developmental Milestones³⁴). The PEDS uses 10 questions to elicit parents' concerns in a range of developmental areas. If any concerns arise, then use of a standardized screener such as the ASQ-3²⁹ or PEDS-Developmental Milestones is recommended.^{34,35} The PEDS-Developmental Milestones provides monitoring of development, using milestone checklists with clear cut-offs for 'typical' or 'concerns' in development.³⁴ There was limited evidence available about the Denver Developmental Screening Test-II; a recent systematic review found it did not meet standards of psychometric evidence.²¹

The revised version of the Parent Report of Children's Abilities is a norm-referenced assessment of cognitive and language development at 2 years of age.^{36,37} It is a parent-completed questionnaire, freely available online. Concurrent validity testing with the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III) successfully predicted cases of cognitive delay and language delay, with receiver operating characteristic curves ranging from 0.83 to 0.97.³⁸ The revised version of the Parent Report of Children's Abilities is recommended by the National Institute for Health and Care Excellence (NICE) for screening 2-year-old children born preterm.³⁹

Multi-domain developmental assessment tools

Multi-domain developmental assessment tools are designed to identify developmental delay. A 2020 systematic review²¹ of multi-domain tools found three developmental assessment tools with good to excellent psychometric properties: Bayley-III,⁴⁰ the Battelle Developmental Inventory, Second Edition (BDI-2),⁴¹ and the Mullen Scales of Early Learning (MSEL)⁴² in order of psychometric quality (Appendix S2). The Bayley-III identifies children and infants who are not meeting developmental milestones across cognitive, language, and motor domains.⁴⁰ The Bayley-III has predictive validity for CP at 4 years: for example, at 2 years, less than 1SD, it has a sensitivity of 0.83 and specificity of 0.94; at less than 2SD, its sensitivity is 0.67 and specificity 1.0.^{43,44} The BDI-2 covers domains of cognition, motor, communication, personal-social, and adaptive skills up to 95 months of age.^{42,45} Both the Bayley-III and the BDI-2 have good psychometric evidence.^{21,46} The Bayley-4 and the BDI-3 are now available; however, no published information of these new versions was available in systematic reviews. The MSEL assesses children aged 0 to 68 months for expressive language, receptive language, gross motor, fine motor, and visual reception.⁴² The MSEL provides an early learning score (all domains except the gross

TABLE 1 Summary of recommended tools by domain and age for screening, assessment, and diagnosis of developmental concerns and disability.

Domain	Screening 0–2 years	Assessment 0–2 years	Screening 2–5 years	Assessment 2–5 years
Multi-domain	<p>Ages & Stages Questionnaires-3^{a,b,29}</p> <p>Ages & Stages Questionnaires - Talking About Raising Aboriginal Kids^{a,c,19}</p> <p>Parents Evaluation of Developmental Status^{a,b,34}</p>	<p>Developmental:</p> <p>Bayley Scales of Infant and Toddler Development, Third Edition⁴⁰</p> <p>Battelle Developmental Inventory, Second Edition⁴¹</p> <p>Mullen Scales of Early Learning⁴²</p> <p>Adaptive functioning:</p> <p>Pediatric Evaluation of Disability Inventory⁵³</p> <p>Pediatric Evaluation of Disability Inventory - Computer-Adaptive Test⁵⁴</p> <p>Pediatric Evaluation of Disability Inventory - Vineland Adaptive Behaviour Scales⁵⁵</p> <p>Adaptive Behavior Assessment system⁵⁶</p>	<p>Ages & Stages Questionnaires-3^{a,b,29}</p> <p>Ages & Stages Questionnaires - Talking About Raising Aboriginal Kids^{a,c,19}</p> <p>Parents Evaluation of Developmental Status^{a,b,34}</p> <p>Parent Report of Children's Abilities - Revised (at 2 years)^{a,b,d,36}</p>	<p>Developmental:</p> <p>Bayley Scales of Infant and Toddler Development, Third Edition⁴⁰</p> <p>Battelle Developmental Inventory, Second Edition⁴¹</p> <p>Mullen Scales of Early Learning⁴²</p> <p>Adaptive functioning:</p> <p>Pediatric Evaluation of Disability Inventory⁵³</p> <p>Pediatric Evaluation of Disability Inventory - Computer-Adaptive Test⁵⁴</p> <p>Vineland Adaptive Behavior Scales⁵⁵</p> <p>Adaptive Behavior Assessment system⁵⁶</p>
Neurological	<p>Predtchl's Qualitative Assessment of General Movements (0–5 months of age) and Motor Optimality Score (3–5 months)^{a,b,60}</p> <p>Alberta Infant Motor Scale⁸⁵</p>	<p>Hammersmith Infant Neurological Examination (2 months–2 years)⁶¹</p> <p>Hand Assessment of Infants (3–12 months)⁶²</p>		<p>Neurological Sensory Motor Developmental Assessment (from 2 years of age)⁶³</p>
Motor		<p>Peabody Developmental Motor Scales, Second Edition⁸⁶</p>	<p>Developmental Coordination Disorder Questionnaire^{a,d,163}</p> <p>Little Developmental Coordination Disorder Questionnaire^{a,164}</p>	<p>Peabody Developmental Motor Scales, Second Edition⁸⁶</p> <p>Movement Assessment Battery for Children, Second Edition (3 years +)⁸⁷</p> <p>Test of Gross Motor Development, Second Edition⁸⁸</p> <p>Bruininks-Oseretsky Test of Motor Proficiency (4+ years)⁸⁹</p>
Fine motor		<p>Hand Assessment of Infants (3–12 months)⁶²</p> <p>Relevant fine-motor domain of multi-domain/ motor assessment tools</p>		<p>Relevant multidomain and motor assessment: e.g. Peabody Developmental Motor Scales, Second Edition;⁸⁶ Motor Assessment Battery for Children, Second Edition;⁸⁷ Mullen Scales of Early Learning;⁴² Battelle Developmental Inventory, Second and Third Edition;⁴¹ Bruininks-Oseretsky Test of Motor Proficiency⁸⁹</p>
Cognition		<p>Bayley Scales of Infant and Toddler Development, Third/Fourth Edition (between 18 months and 40 months)⁴⁰</p> <p>Griffiths Mental Development Scales⁵²</p>		<p>Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition¹⁰⁵</p> <p>Stanford-Binet Intelligence Scales¹⁷⁶</p> <p>Differential Ability Scales-II¹⁷⁷</p> <p>Developmental Neuropsychological Assessment¹⁷⁶</p> <p>Revised Leiter International Performance Scale¹⁰⁶</p>

TABLE 1 (Continued)

Domain	Screening 0–2 years	Assessment 0–2 years	Screening 2–5 years	Assessment 2–5 years
Communication/ language	Communication and Symbolic Behavior Scales (Infant-Toddler) (6–24 months) ^{a,108} Hearing and Talking Scale ^{b,c,d,17} Parents' Evaluation of Listening and Understanding Measure ^{b,c,d,16}	Preschool Language Scales, Fifth Edition ¹¹⁴	Communication and Symbolic Behavior Scales (Infant-Toddler) (6–72 months for children with significant delay) ^{a,108} Hearing and Talking Scale ^{b,c,d,17} Parents' Evaluation of Listening and Understanding Measure ^{b,c,d,16}	Speech: Diagnostic Evaluation of Articulation and Phonology ¹¹³ Dynamic Evaluation of Motor Speech Skill ¹¹¹ Language: Preschool Language Scales, Fifth Edition ¹¹⁴ Clinical Evaluation of Language Fundamentals Preschool, Second Edition (3–7 years) ¹¹⁵ Clinical Evaluation of Language Fundamentals, Fifth Edition (5+ years) ¹¹⁵
Social–emotional	Brief Infant/Toddler Social and Emotional Assessment ¹¹⁶ The Child Behaviour Checklist 1.5/5 ^{b,123} Ages & Stages Questionnaires: Social–Emotional, Second Edition ¹¹⁸	Infant Toddler Social and Emotional Assessment (1–3 years) ¹¹⁷ Greenspan Social–Emotional Growth Chart (0–3.5 years) ¹²⁴	Strengths & Difficulties Questionnaire ^{a,b,d,119} Ages & Stages Questionnaires: Social–Emotional, Second Edition ¹¹⁸	The Behaviour Assessment System for Children (2 years +) ¹²² The Child Behaviour Checklist ¹²³ Greenspan Social–Emotional Growth Chart (0–3.5 years) ¹²⁴ Infant Toddler Social and Emotional Assessment (1–3 years) ¹¹⁷ The Indigenous Child-Initiated Pretend Play Assessment with its component of the Indigenous Play Partner Scale ^{c,18}
Sensory processing		Test of Sensory Function in Infants – 4 to 18 months ⁹⁶ Infant/Toddler Sensory Profile 2 ¹⁷⁷		Participation and Sensory Environment Questionnaire–Home (2–7 years) ^{178,179} Child Sensory Profile-2 (3–14 years) ⁹⁸ Sensory Processing Measure–Preschool (2–5 years) ¹⁸⁰
Diagnostic group	Screening 0–2 years	Assessment 0–2 years	Screening 2–5 years	Assessment 2–5 years
Cerebral palsy	Prechtl's Qualitative Assessment of General Movements (0–5 months) ^{a,b,60}	Hammersmith Infant Neurological Examination (2–24 months) ⁶¹ Developmental Assessment of Young Children ¹²⁷ Hand Assessment of Infants ⁶²	Gross Motor Function Classification System ¹²⁹ Manual Ability Classification System ¹³⁰ Communication Function Classification System ¹³¹	Gross Motor Function Measure ¹³²
Attention-deficit/hyperactivity disorder	Child Behaviour Checklist 1.5/5 ^{a,b,123} Strengths and Difficulties Questionnaire ^{a,b,d,119}	Behavior Assessment System for Children ¹²² Behavior Rating Index of Executive Function–Preschool (BRIEF-P) ¹⁵⁷ Brown Attention Deficit Disorder Symptom Assessment Scale ¹⁵⁸	Strengths and Difficulties Questionnaire ^{a,b,d,119} Child Behaviour Checklist 1.5/5 ^{a,b,123}	Measures of attention and behaviour: Behavior Assessment System for Children, Third Edition ¹²² Conners Early Childhood Ratings Scales ¹⁵⁹ Focus on the executive function: BRIEF-P ¹⁵⁷ Brown Attention Deficit Disorder Symptom Assessment Scale ¹⁵⁸

(Continues)

TABLE 1 (Continued)

Domain	Screening 0–2 years	Assessment 0–2 years	Screening 2–5 years	Assessment 2–5 years
Autism spectrum disorder	Revised Social Attention and Communication Surveillance ^{a,b,135} ASDetect smartphone app ^{a,b,d,137} Revised Modified Checklist for Autism in Toddlers with Follow-up ^{a,b,d,136}	The Autism Diagnostic Observational Schedule, Second Edition (12 months +) ¹⁵⁰ Childhood Autism Rating Scale ¹⁵¹ Autism Observation Scale for Infants (6–36 months) ¹⁵³	Social Attention and Communication Surveillance Preschool ^{a,b,135} Revised Modified Checklist for Autism in Toddlers with Follow-up ^{a,b,d,136}	Autism Diagnostic Observational Schedule, Second Edition ¹⁵⁰ Childhood Autism Rating Scale ¹⁵¹ Revised Autism Diagnostic Interview ¹⁵²
Developmental coordination disorder			Developmental Coordination Disorder Questionnaire ^{a,b,d,163} Little Developmental Coordination Disorder Questionnaire ^{a,b,165}	Movement Assessment Battery for Children, Second Edition ⁸⁷
Fetal alcohol spectrum disorder	Routinely gathering information on prenatal alcohol exposure when taking developmental history ^{11,166}	For a diagnosis of fetal alcohol spectrum disorder three criteria must be considered: (1) prenatal alcohol use (AUDIT-C) ¹⁶⁷ and other exposures, (2) neurodevelopmental impairment, and (3) sentinel facial features	Routinely gathering information on prenatal alcohol exposure when taking developmental history ^{11,166}	For a diagnosis of fetal alcohol spectrum disorder, three criteria must be considered: (1) prenatal alcohol use (AUDIT-C) ¹⁶⁷ and other exposures, (2) neurodevelopmental impairment, and (3) sentinel facial features

^aAvailable for use by all health professionals (general practitioners, nurses, Indigenous health workers, allied health professionals).

^bTelehealth-friendly.

^cCulturally adapted for Australian Aboriginal and Torres Strait Islander children.

^dFreely available (no cost).

motor scale) and has validity for cognitive assessment.^{47–49} The MSEL has been used for the prospective assessment of infant siblings of children with autism.^{50,51} The MSEL has separate scales for visual reception (non-verbal abilities) and receptive and expressive language;⁴² it detects asymmetry in developing abilities. The revised version of the Griffiths Mental Developmental Scales⁵² was found to have unsatisfactory psychometric properties.²¹

Multi-domain functional/adaptive assessment tools

Multi-domain functional/adaptive assessment tools are used to measure functional performance and adaptive behaviour across several domains. The Pediatric Evaluation of Disability Inventory (PEDI),⁵³ Pediatric Evaluation of Disability Inventory - Computer-Adaptive Test (PEDI-CAT),⁵⁴ the Vineland Adaptive Behavior Scales (VABS-2/3),⁵⁵ and the Adaptive Behavior Assessment System, Third Edition (ABAS-3)⁵⁶ are standardized multi-domain assessments based on large normative samples, with the primary purpose of detecting functional delay. The PEDI is a paper questionnaire of self-care, mobility, and social function for children aged 6 months to 7 years 6 months.⁵³ The PEDI-CAT, which is based on the PEDI, was refined through Rasch modelling, with items that extended the age range up to 20 years.⁵⁴ It uses computerized adaptive testing to minimize the time for completion. Both the VABS and the ABAS cover a wide age range (birth to 90/89 years respectively) using parent/caregiver questionnaires. The VABS measures four domains (daily living skills, motor skills, communication, and socialization) which provide an adaptive behaviour composite (for ages from birth–6 years 11 months).⁵⁵ The ABAS-3 measures adaptive behaviour skills across social, practical, and conceptual domains.^{56,57}

The WeeFIM (child version of Functional Independence Measure) is a tool that measures the severity of disability in two main domains of motor and cognition.^{58,59} It measures the type and amount of assistance required for a person with a disability to perform basic life activities effectively. It is often used to measure outcomes of rehabilitation following acquired/traumatic brain injury or burns. Training is required to be certified to use the WeeFIM.

Neurological domain tools

Of the 15 neurological tools identified for children aged birth to 5 years, there is one screening tool and three assessment tools recommended to identify infants with a higher chance of adverse neurodevelopmental outcomes. These tools are recommended at specific timepoints for children with a history of neurological risk factors (e.g. born preterm, low birthweight, stormy neonatal period), or strong parental concern about development: (1) Prechtl's General Movements Assessment (GMA) (0–5 months of age);⁶⁰ (2) Hammersmith Infant

Neurological Examination (HINE) (3 months–2 years of age);⁶¹ (3) the Hand Assessment for Infants (HAI) (3–12 months of age),⁶² and (4) the Neurological, Sensory, Motor, Development Assessment (NSMDA) (>2 years of age).⁶³

Neurological screening tools for 0 to 5 months of age

Prechtl's General Movements Assessment (GMA) is the screening tool with the best prediction of high risk of CP (98% sensitivity, specificity)^{64–68} and demonstrates strong ability to differentiate between infants who are typically developing and those with a higher chance of motor or cognitive delays.⁶⁹ The GMA uses a simple 3- to 5-minute video of an infant's spontaneous movements to screen infants aged from birth to 5 months corrected age.⁷⁰ Scoring is performed by specially trained assessors. Prediction of outcomes is most accurate for GMA trajectories across time points (writhing and fidgety periods).^{60,69} Emerging evidence supports the concurrent analysis using the revised version of the Motor Optimality Score⁷¹ to provide predictive information for risk of non-CP adverse neurodevelopmental outcomes^{72,73} and functional severity of CP.^{71,72} The GMA has demonstrated feasibility both cross-culturally and in low-resource settings.^{71,74}

The Hammersmith Neonatal Neurological Examination, the Neonatal Network Neurosensory Scale, and the Amiel-Tison Neurological Assessment at Term are not included in our recommendations as they are specifically used by specialist clinicians working in neonatal settings and have a very small time-window for use.^{66,75,76,77} These tools do not have evidence for use through telehealth and provide less predictive information than the GMA and HINE.

Neurological assessment tools for 3 months to 2 years corrected age

Infants who are screened to be at higher risk of adverse neurodevelopmental outcomes on the GMA or those who present with neurological concerns after 5 months corrected age can be referred for assessment using the HINE.⁶¹ The HINE can be used to assess a child up until 2 years corrected age. Compared with other neurological assessments, the HINE is inexpensive, accessible, and has excellent predictive accuracy for CP⁷⁸ and severe motor⁶⁹ and cognitive delay.⁷⁹ Scores on the HINE provide further information about CP topography and prediction of gross motor skill attainment.⁷⁸ Infants who present with five or more asymmetries on the HINE have a higher chance of unilateral CP,⁸⁰ and warrant further investigation with a tool such as the HAI. A 2017 systematic review recommended the Test of Infant Motor Performance for early diagnosis of CP when the GMA and magnetic resonance imaging are unavailable;¹⁵ however, the HINE's broader age range and transdiagnostic utility for predicting outcomes other than CP support its use in preference to the Test of Infant Motor Performance.^{69,79,81}

The HAI⁶² is recommended for early detection of unilateral CP.⁸² The HAI has strong psychometrics (Appendix S2). Accuracy of diagnosis of unilateral CP improves after 5 months of age.^{80,82} Interrater and test-retest reliability are high (ICC 0.92–0.99).⁸³ The HAI provides age-normative values for unimanual and bimanual activity.⁸⁴

Neurological assessment tools for >2 years

Children who present after 2 years with neurological risk factors or strong parental concerns should be referred for assessment. The recommended tool at this age is the NSMDA a combination of observation and hands-on assessment. The NSMDA discriminates between children who are developing typically and those with mild, moderate, or severe delays predictive of neurodevelopmental outcomes.⁴³ Predictive accuracy improves with increasing age.⁴³

Motor domain tools

Fourteen tools were identified and evaluated that assessed motor outcomes in children aged birth to 5 years 11 months. Eight tools evaluated fine and gross motor outcomes and six assessed gross motor development only. Recommendations include one screening tool (the Alberta Infant Motor Scale [AIMS]⁸⁵) and four standardized assessment tools (the Peabody Developmental Motor Scales, Second Edition [PDMS-2];⁸⁶ Motor Assessment Battery for Children, Second Edition [MABC-2];⁸⁷ the Test of Gross Motor Development, Second Edition [TGMD-2];⁸⁸ and the Bruininks-Oseretsky Test of Motor Proficiency [BOT-2]⁸⁹) to identify infants with a higher chance of ongoing motor delays.⁴³

Motor domain 0 to 3 years corrected age

Motor screening < 18 months

The AIMS⁸⁵ is an observational, norm-referenced tool used to screen infant gross motor development until 18 months corrected age. It discriminates between infants who are developing typically and those with delays, and has good predictive accuracy at 4 months, 8 months, and 12 months to predict motor delays at 18 months.⁴⁴ The AIMS is a reliable, cost-effective tool which is quick and easy to administer. Owing to its observational nature, it is easily incorporated into routine clinical appointments and can be completed through telehealth.⁴⁴ The AIMS demonstrates good concurrent validity with the motor quotients on the PDMS-2 and the Bayley Scales of Infant Development, Second Edition at 12 months.⁹⁰

Motor assessment ≤ 3 years

The PDMS-2⁸⁶ assesses a child's gross and fine-motor skills to discriminate between those developing typically and

those with mild to severe delays on the basis of a normative sample. The PDMS-2 has very weak predictive validity at 2 years when completed before 8 months corrected age;⁴⁴ however, at 3 years it demonstrates good predictive validity for identifying probable DCD.⁹¹ The PDMS-2 has a lengthy administration time and cannot be administered through telehealth.

Motor domain 3 to 5 years of age

Three tools are recommended⁴³ to assess motor development of children aged 3 to 5 years: the MABC-2,⁸⁷ the TGMD-2,⁸⁸ and the BOT-2,⁸⁹ ranked in order of predictive validity and clinical utility (Appendix S2). The MABC-2 is a discriminative and predictive tool that assesses both fine and gross motor components and is often used for diagnostic criteria of DCD (MABC < 15th centile).^{14,92} It has predictive ability and stability in the preterm population; at 4 years it predicts moderate to severe motor delay at 7 to 9 years.⁹¹ The MABC-2 has a short administration time and is validated for use in telehealth up to 16 years.⁹³ The MABC-3 is now available; however, no published information of this new version was available in systematic reviews. The TGMD-2 is a quick and feasible discriminative assessment for gross motor development in children from 3 years. It demonstrates good test-retest and interrater reliability, but lacks the ability to identify ongoing motor delays or DCD.^{43,91} The TGMD-2 has a shorter administration time (15–20 minutes) than the MABC-2 (20–40 minutes), but does not include a fine-motor component. The BOT-2 is a lengthy assessment, with weaker predictive validity for later motor delay.⁸⁹ A short form of the BOT-2 is available which is not feasible through telehealth.

Fine-motor skills assessment in the context of general motor development

Two screening tools contain fine-motor components: the ASQ-3 and the Denver Developmental Screening Test-II (see the section on 'Multi-domain tools'). Three standardized motor assessment tools contain a fine-motor domain (PDMS-2, MABC-2, BOT-2) and three standardized multi-domain assessment tools contain a fine-motor domain (Bayley-III, BDI-2, MSEL). These are all normative tests administered in a standardized manner to discriminate whether motor skill development is typical or delayed. The fine-motor component of the Bayley-III before 6 months has best predictive accuracy of all the Bayley-III domains for outcomes at 2 years of age when tracking the development of high-risk infants born preterm.⁶⁹ A study of concurrent validity of Bayley-III and PDMS-2 in children born preterm at 18 months found 97% agreement on fine-motor skills.⁹⁴

Sensory domain tools

Sensory dysregulation refers to atypical responses to sensory stimuli, including inability to organize and regulate sensory stimuli (including touch, hearing, vision, smell, movement, and balance) and is often observed in children with neurodevelopmental disorders.⁹⁵ Sensory measures are generally questionnaires completed by a parent/caregiver.

Sensory domain <2 years corrected age

A systematic review of sensory measures for children aged younger than 2 years recommended the Sensory Profile 2.⁹⁶ The Sensory Profile 2 is a questionnaire for use in infants up to 6 months of age and a toddler version for 7 to 35 months of age which can be used through telehealth. The Test of Sensory Function in Infants is a performance-based assessment that measures sensory systems (vision, tactile pressure, vestibular) for children aged 4 to 18 months of age (Appendix S2).⁹⁶

Sensory domain >2 years

A recent systematic review⁹⁵ recommended the Participation and Sensory Environment Questionnaire-Home (PSEQ-H) to evaluate sensory dysregulation between 2 years and 7 years.⁹⁷ The PSEQ-H has moderate to high quality for structural validity, hypothesis testing for construct validity, internal consistency, reliability, and measurement error. The PSEQ-H is freely available online.⁹⁵ Both the Child Sensory Profile-2^{95,98} and the Sensory Processing Measure - Preschool were noted to have sufficient ratings of discriminant validity with moderate levels of evidence.⁹⁵ The Child Sensory Profile-2 was designed for children aged 3 years to 14 years 11 months to discriminate between children who had a diagnosis of autism, ADHD, or other neurodevelopmental disorders.⁹⁵ It lacks evidence for content validity. The SPM-P is used for 2- to 5-year-old children.

Cognitive domain tools

Intellectual disability is characterized by deficits in general mental abilities, such as problem-solving, planning, reasoning, and judgement.¹ Three key criteria are used to establish a diagnosis of intellectual disability: (1) deficits in intellectual functions, confirmed by both clinical assessment and individualized, standardized intelligence testing; (2) deficits in adaptive functioning that limit functioning in one or more activities of daily life, across multiple environments; and (3) onset of intellectual and adaptive deficits during development.¹ Global developmental delay is a diagnosis used for children younger than 5 years.¹ Children who fail

to meet expected developmental milestones in several areas of functioning, or who are unable to undergo systematic assessments of intellectual functioning, or are too young to participate in standardized testing, may receive a diagnosis of global developmental delay.¹

A 2016 meta-analysis of early developmental assessment tools used to assess cognition in children born very preterm or with very low birthweight found the Bayley and the revised version of the Griffiths Mental Developmental Scales were the most commonly used cognitive tests between 18 months and 40 months of age.⁹⁹ Early assessments (both Bayley and the revised version of the Griffiths Mental Developmental Scales) were fairly accurate in predicting the absence of school-age cognitive deficits (high negative predictive value, range 47.8–95.5%); however, prediction of any cognitive impairment was weak, with a pooled sensitivity 55% (95% CI 45.7–63.9%) and specificity 84.1% (77.5–89.1%).⁹⁹ Meta-analysis suggested half of children who may have cognitive difficulties at school-age were classified as having typical neurodevelopmental function at 1 to 3 years of age.⁹⁹ This highlights the importance of repeated assessments for children whose families/carers have concerns about their child's development.

Comprehensive measures of cognition were noted as the Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV), the Stanford-Binet Intelligence Scales, Fifth Edition (SB-5), the Differential Ability Scales, Second Edition (DAS-II), Developmental NEuroPSYchological Assessment (NEPSY-II), and the revised version of the Leiter International Performance Scale.^{22,46,100,101,102,103,104} These may all be used in primary health settings by appropriately qualified professionals. The WPPSI-IV is a measure of cognitive development for preschoolers and young children (2 years 6 months–3 years 11 months; 4 years–7 years 7 months). Validity studies of the WPPSI-IV with other measures for targeted populations (e.g. children with ADHD, autism, intellectual disability) are published in the manual.¹⁰⁵ The WPPSI-III correlated highly with the Bayley Scales of Infant Development, Second Edition and DAS (>0.8) in studies of children with prenatal exposure to anti-depressants.²² The WPPSI-IV has been standardized for children aged 2 years 6 months to 7 years 7 months.¹⁰⁵ The SB-5 (Early SB5) comprehensively measures cognitive ability of children aged 2 years to 7 years 3 months. The Full-scale IQ consists of 10 subtests, with a verbal and non-verbal subtest for five factors: fluid reasoning, knowledge (including vocabulary), quantitative reasoning, visual–spatial processing, and working memory. The Stanford Binet Intelligence Scale has good test–retest and interrater reliability ($\alpha=0.83$).²² The DAS-II measures general cognitive abilities for children aged 2 years 6 months to 17 years 11 months.^{22,102} The subtests for the Early Years battery differ from that of the School-Age battery. The Early Years battery (2 years 6 months–3 years 5 months) contains only four subtests covering verbal and non-verbal ability, while the upper level (3 years 6 months–6 years 11 months) contains six subtests which cover verbal, non-verbal, and spatial ability clusters. The DAS-II allows for quick administration.¹⁰²

The NEPSY-II is a comprehensive measure of neuropsychological development in children aged 3 to 16 years.^{22,101} It includes assessment domains of attention and executive functioning, language, memory and learning, sensorimotor, social perception, and visuospatial processing.

The revised version of the Leiter International Performance Scale is a non-verbal assessment for cognitive abilities for children and young people aged between 2 years and 20 years.¹⁰⁶ It is suitable for children who have a communication impairment but typical visual function.^{103,104} It measures cognitive abilities including fluid reasoning, visuospatial ability, short-term memory, long-term retrieval, processing speed, and general knowledge.^{103,104}

Diagnosis of an intellectual disability requires assessment of adaptive functioning. Measures of adaptive functioning commonly used to inform diagnosis of intellectual developmental disorder include multi-domain tools such as the Vineland-3,⁵⁵ ABAS-3,⁵⁶ and PEDI-CAT.⁵⁴ The Vineland-3 scales measure three broad domains of adaptive functioning: communication, daily living skills, and socialization. Telehealth assessment of cognitive functions is supported.¹⁰⁷

Communication/language domain tools

There are many skills that contribute to effective communication (speech, receptive and expressive language, social communication, fluency, voice, pre-literacy). In the preschool years, evaluation should include screening and/assessment of speech, and receptive and expressive language. Clinical practice guidelines¹² have identified tools available to support assessment of children's communication.¹² The tools described in this scoping review cover multiple communication domains in a single measure. Tools that assessed only a single communicative domain were not recommended owing to their limited application, despite some having strong psychometrics (e.g. the Language Use Inventory, the Peabody Picture Vocabulary Test, the Renfrew Bus Story).

Communication domain screening tools

The Communication and Symbolic Behavior Scales (Infant-Toddler; CSBS) is a parent-completed screening questionnaire of early verbal and non-verbal communication for children aged 6 to 24 months (up to 72 months for children with significantly delayed communication).¹⁰⁸ It covers a range of communication domains including functional ability to convey different communicative intents, use of gestures, speech, receptive language, and expressive vocabulary. Raw scores can be converted to standard scores and centile ranks to identify children who require further assessment. At 12 to 17 months, the CSBS screener has fair predictive validity with language evaluation on the Preschool Language Scales, Third Edition, or the MSEL at 2 to 3 years.¹⁰⁹ The CSBS (Infant-Toddler) was recommended as a screening tool for early signs of autism at 12 months owing to its high

predictive value (sensitivity=0.93, specificity=0.89) and clinical utility.¹¹⁰ The Hearing and Talking Scale (HATS)¹⁸ is a screening tool for detecting communication problems in young Indigenous Australian children aged 4 to 30 months. Validity of the HATS was performed with the ASQ-TRAK and the Expressive Vocabulary Test, Second Edition.¹⁷ Similarly, the Parents' Evaluation of Listening and Understanding Measure (PLUM) is a co-designed screening tool for screening functional auditory performance of Indigenous Australian children.¹⁶ The PLUM can identify children at risk of hearing and listening problems.¹⁶

Standardized speech assessment tools

The Dynamic Evaluation of Motor Speech Skill is a criterion-referenced assessment designed to differentially diagnose severe speech sound disorders in children older than 3 years, which must be administered by a speech pathologist.¹¹¹ It has fair predictive validity for identifying childhood apraxia of speech (sensitivity=0.9, specificity=0.7)¹¹² and strong test-retest (ICC=0.82), and intrarater (ICC=0.92) and interrater reliability (ICC=0.98).^{111,112} The Diagnostic Evaluation of Articulation and Phonology differentially assesses speech in children aged over 3 years (articulation, phonology, and oral motor), and includes both screening and in-depth assessment. Although there are limited psychometrics for the Diagnostic Evaluation of Articulation and Phonology¹¹³ it is a widely used assessment of speech.¹²

Standardized language assessment tools

The Preschool Language Scales, Fifth Edition¹¹⁴ is a norm-referenced assessment of receptive and expressive language for children from birth to 7 years 11 months. It has good predictive validity (sensitivity=0.91, specificity=0.78) for children at risk of language disorder.¹⁰⁹ It has strong content and construct validity.¹⁰⁹ A screening version (domains of language, articulation, connected speech, social/interpersonal communication skills, fluency, and voice) was designed for use by early educators.

The Clinical Evaluation of Language Fundamentals Preschool, Second Edition¹¹⁵ and Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF-5) are widely used norm-referenced assessments of receptive and expressive language (including vocabulary, sentences, and functional communication) for children aged 3 years 6 months to 6 years 11 months, and 5 years to 21 years respectively.¹² Psychometrics of the Clinical Evaluation of Language Fundamentals Preschool, Second Edition report poor to fair predictive validity (sensitivity=0.42, specificity=0.99) for identifying language disorders (using cut-point -2SD), and fair predictive validity for 'at risk of language disorder' (using cut-point -1.25SD; sensitivity=0.64, specificity=0.93).¹⁰⁹ The CELF-5 has good predictive validity for identifying children 'at risk of language disorders' (-1.5SD; sensitivity=0.85,

specificity=0.99), but poor predictive validity for confirmed language disorder (-2SD; sensitivity=0.57, specificity=1.0). It has strong content and construct validity and reliability, with limited psychometrics across other domains.¹⁰⁹

Social-emotional domain

Social-emotional development in children from birth to 5 years 11 months encompasses the attainment of expected developmental milestones in social and emotional domains as well as childhood mental health. Typical socio-emotional development in this age group includes joint attention, social interest, pretend play, developing theory of mind (understanding perspectives of other people), developing emotional regulation, and the emergence of social skills such as taking turns or sharing. Some assessments in the social-emotional domain focus on the attainment of developmental milestones while others focus on early mental health. In early childhood, mental health is understood in terms of externalizing and internalizing behaviour. Externalizing behaviours are directed outwards and may include aggression, temper tantrums, and non-compliance. Internalizing behaviours are directed inwards and may include feelings of sadness or anxiety, social withdrawal, or physiological symptoms such as stomach aches.

There is overlap between both the developmental and the mental health aspects of the social-emotional domain and the ADHD and autism-related symptomology which are reported under the specific diagnostic groups. Assessments are recommended on the basis of their psychometric properties and clinical utility. Some assessments have age ranges beyond 5 years 11 months of age (the Ages & Stages Questionnaires: Social-Emotional, Second Edition [ASQ-SE-2], and Strengths and Difficulties Questionnaire [SDQ]), so the psychometrics reported may be an underestimation, reflective of limited studies in children of this age. Assessment choice must consider the focus of the assessment: developmental attainment or mental health or both. The child's age must be taken into account, and whether it is desirable to choose an assessment that can be repeated later.

Social-emotional screening tools

The Brief Infant Toddler Social and Emotional Assessment (BITSEA),¹¹⁶ the SDQ, and the ASQ-SE-2 are recommended screening tools for social-emotional outcomes. The BITSEA is a brief screener version of the Infant Toddler Social and Emotional Assessment (ITSEA).^{116,117} It is a short caregiver-report measure for children 1 to 3 years of age. It gives total scores for social-emotional problems (including internalizing and externalizing behaviour) and social-emotional competence. A systematic review²⁴ identified good to excellent internal consistency and agreement with the Child Behaviour Checklist (CBCL) (0.51–0.79), and prediction of CBCL and ITSEA scores 1 year later (Appendix S2). The

ASQ-SE-2 is a caregiver questionnaire for children from 1 month to 6 years to assess social-emotional competency across self-regulation, compliance, social communication, adaptive functioning, autonomy, affect, and social interaction.¹¹⁸ The ASQ-SE-2 has excellent internal consistency (0.84) and interrater reliability (0.91), high agreement with other measures (0.81–0.95), and good to very good predictive validity (sensitivity=0.81, specificity=0.84).²⁴ The SDQ¹¹⁹ is a brief caregiver-report screening tool of internalizing and externalizing behaviour as well as prosocial behaviours and peer relations for children from 2 to 4 years, and from 4 to 17 years. It is a 25-item measure of child behaviour and adjustment. The SDQ has adequate internal reliability ($\alpha=0.73$) and test-retest reliability ($r=0.62$) as well as discriminant and concurrent validity.¹²⁰ The SDQ is freely available, with no restrictions on who can administer it.

Social-emotional assessment tools

Tools recommended for assessment of social-emotional outcomes in early childhood are the Behaviour Assessment System for Children (BASC), CBCL, Greenspan Social-Emotional Growth Chart, and ITSEA.^{10,24,121}

The BASC is a widely used caregiver report for children 2 years and older.¹²² The third edition, the BASC-3, measures externalizing and internalizing symptoms as well as social well-being with four composite scales for preschool children: internalizing problems, externalizing problems, adaptive skills, and the behavioural symptoms index. The BASC-2 has fair internal consistency and fair to good construct validity.¹²¹ The CBCL is a parent or caregiver-report questionnaire, and the preschool version is appropriate for children aged 1 year 6 months to 5 years.¹²³ It is a mental health measure, assessing both externalizing and internalizing behaviour. The CBCL is part of the Achenbach System of Empirically Based Assessment. A systematic review found the CBCL correctly classified 84% of children with emotional and behavioural problems.²⁴ The Greenspan Social-Emotional Growth Chart is a caregiver-completed questionnaire used from birth to 3 years 6 months.¹²⁴ It is a developmental assessment, focused on the child's mastery of developmental tasks in the social-emotional domains. It has excellent internal consistency ($\alpha=0.83$ –0.94), good construct validity as measured by agreement with the Bayley-III, and predictive validity (sensitivity=0.87, specificity=0.90) as a screening tool for ASD.^{24,125} The ITSEA is a caregiver-report measure for children aged 1 to 3 years.¹¹⁷ It is both a mental health and developmental assessment covering externalizing and internalizing behaviour as well as competence across play, empathy, and prosocial peer relations. The ITSEA has good to excellent internal consistency (externalizing 0.66, internalizing 0.85, competence 0.56–0.79) and agreement with the CBCL (range 0.41–0.60).²⁴

The Indigenous Child-Initiated Pretend Play Assessment with its component of the Indigenous Play Partner Scale is an assessment of play skills and social behaviours which

has been validated for Australian Aboriginal children.^{18,126} Further information on this tool is given in Appendix S4.

Tools specific to diagnostic groups

Cerebral palsy

CP is diagnosed on the basis of a combination of clinical and neurological signs.⁴ The International Clinical Practice Guidelines for the early, accurate diagnosis of CP found that early diagnosis begins with medical history then use of standardized neurological and motor assessments together with neuroimaging.¹⁵ Before 5 months corrected age, the most predictive tools are the GMA (sensitivity=0.98)⁶⁰ and the HINE (sensitivity=0.90).⁶¹ After 5 months of age, the most predictive tools were the HINE (sensitivity=0.90)⁶¹ and the Developmental Assessment of Young Children (sensitivity=0.83).^{15,127} Any infant unable to sit independently by 9 months, presenting with hand function asymmetry, or unable to take weight through the plantar surface of the feet, should receive an investigation for CP.¹⁵ The functional consequences of CP are described using classifications systems: the Gross Motor Function Classification System,^{128,129} the Manual Ability Classification System,¹³⁰ and the Communication Function Classification System.¹³¹ Each comprises five ordinal levels to discriminate meaningfully on a child's ability. Level I represents the highest level of ability and level V represents the lowest. The Gross Motor Function Measure (GMFM)¹³² is a condition-specific measure of the capacity for gross motor activity in children aged 5 months to 16 years with a diagnosis of CP. The extensively validated versions include the original GMFM-88 item, and the GMFM-66 item which was developed and validated by Rasch analysis.^{35,58} A recent systematic review⁵⁸ identified that the GMFM-66 and GMFM-88 have sufficient responsiveness. Ceiling effects in children with high motor ability occur as the most difficult items are at a level that a typically developing 5-year-old can achieve. The GMFM-66 can discriminate between children with CP who have different levels of motor ability.^{133,134}

Autism

National guidelines for screening and assessing autism are recommended for clinicians working with children at risk of autism.¹³ The guidelines highlight the importance of comprehensive multidisciplinary assessment for autism, gathering information through a combination of observation, and standardized and non-standardized tools.

Screening tools for autism

There are several screening tools that can be used to identify children under the age of 5 years showing early signs of autism: the revised version of the Social Attention

and Communication Surveillance (SACS-R); the Social Attention and Communication Surveillance Preschool (SACS-PR);¹³⁵ the revised version of the Modified Checklist for Autism in Toddlers, with Follow-up (M-CHAT-R/F);¹³⁶ and a freely available smartphone app for parents, ASDetect.^{135,137,138,139,140,141,142,143} The SACS-R is recommended as the most effective screening tool from 11 to 30 months, and the SACS-PR for children aged 3 years 6 months to 5 years.¹³⁵ The SACS-R is an observational checklist completed by trained clinicians which takes approximately 10 to 15 minutes to administer. The SACS-R has strong predictive validity.¹³⁵ The M-CHAT-R/F¹³⁶ is a freely available clinician-led parent questionnaire that can be used to screen children 16 to 30 months.¹⁴¹⁻¹⁴³ Although the sensitivity and specificity of this tool are not as reliable as those tools above, the M-CHAT-R/F is readily accessible and does not require any training or costs. A meta-analysis ($n = 49\,841$) found the M-CHAT-R/F had a pooled sensitivity of 82.6% and specificity of 45.7%, with more effective predictive value when used in high-risk populations.¹⁴⁴ The M-CHAT-R/F had a sensitivity of 73% and specificity of 89% for diagnosis of autism when a cut-off score of 3 was used with a sample of 16 071 children.¹⁴⁵ Children identified as being at risk of autism using these screening tools should be referred for further diagnostic assessment and encouraged to access early intervention and family supports. ASDetect is a freely available smartphone app designed for parents and caregivers on the basis of the SACS-R.¹³⁷ The app provides video examples to help parents/caregivers answer questions related to their child and takes 20 to 30 minutes to administer. If a child is identified as having a high likelihood of autism, parents can seek referral from their medical practitioner for further assessment and support.

Diagnostic tools for autism

Many of the recommended diagnostic assessment tools require significant training and expertise. It is best practice that diagnostic assessments are done by a multidisciplinary team rather than an individual practitioner.^{13,146,147,148,149} Diagnostic tools that most accurately diagnose autism in preschool children identified in a Cochrane review¹⁴⁹ were the Autism Diagnostic Observational Schedule, Second Edition (ADOS-2),¹⁵⁰ the Childhood Autism Rating Scale (CARS),¹⁵¹ and the revised version of the Autism Diagnostic Interview (ADI-R).¹⁵² The ADOS is best for not missing children who have autism and is similar to the CARS and ADI-R in not falsely diagnosing autism. Diagnostic tools should be combined with observations and information from a range of sources by specialized clinicians as recommended in guidelines.¹³

The ADOS-2 is the recommended standardized assessment tool for autism diagnosis used with children aged 12 months and above.^{147,149} This is a standardized, behavioural observation and play assessment that takes 40 to 60 minutes to administer. Meta-analysis of the ADOS-2 found a sensitivity of 0.94 (95% CI 0.89–0.97) and specificity

of 0.80 (95% CI 0.68–0.88).¹⁴⁹ Overdiagnosis is likely if the ADOS is used in populations with lower prevalences of autism.¹⁴⁹ The CARS-2 combines an interview with the parents/caregivers and observation of the child. It is used with children 2 years and older.¹⁵¹ The CARS-2 has a combined sensitivity of 0.8 (95% CI 0.61–0.91) and specificity of 0.88 (95% CI 0.64–0.96).¹⁴⁹ The ADI-R is a diagnostic interview that can be used with children 2 years and older.¹⁵² The Autism Observation Scale for Infants¹⁵³ is the most reliable tool for detecting autism in high-risk populations under 36 months (e.g. children with an older sibling who has a diagnosis of autism; sensitivity=0.84, specificity=0.98).¹⁴⁵ The Autism Observation Scale for Infants is an observational checklist completed by clinicians, which takes 15 to 20 minutes.¹⁵³ This tool is only used in research and has not been standardized as a clinical screening or diagnostic tool.

ADHD

ADHD is a neurodevelopmental disorder in which persistent inattention and/or hyperactivity/impulsivity affect functioning.¹ It begins in childhood, and features must be present before 12 years. For the inattentive presentation, at least six symptoms associated with inattention must be present such as difficulty with sustaining attention in tasks, not following through on instructions, difficulty with organizing tasks, or not paying attention to details and making careless mistakes. Diagnosis of the hyperactive/impulsive presentation requires at least five symptoms, such as fidgetiness, excessive talking, blurting out responses, restlessness, and difficulty sitting still.¹ For all presentations, symptoms must be present across two or more settings. Symptoms must interfere with or reduce quality of functioning.¹ Recent clinical practice guidelines on the diagnosis of ADHD¹⁰ highlight the importance of comprehensive assessment. Therefore, all the tools discussed are best considered as screening measures. Whereas some are multi-domain scales that include measures of attention and behaviour (e.g. the BASC-3, the Conners Early Childhood Assessment [Conners EC], the CBCL, and the SDQ), others focus on the executive function challenges associated with ADHD (Behavior Rating Index of Executive Function–Preschool [BRIEF-P], Brown Attention Deficit Disorder Symptom Assessment Scale [BADDS]/Brown Executive Function/Attention Scales).

Screening tools for ADHD

The following tools have potential utility for screening and/or assessment for ADHD in children up to 5 years 11 months:^{10,154,155,156} the BASC-3,¹²² BRIEF-P,¹⁵⁷ BADDS,¹⁵⁸ CBCL,¹⁵⁴ Conners EC,¹⁵⁹ and SDQ.¹¹⁹

The BASC-3 is a tool used for social–emotional assessment and ADHD. It measures children's adaptive and problem behaviours and emotional difficulties at home and in community settings using parent and teacher questionnaires.¹²² Scores provide clinical and adaptive scales including attention

problems and hyperactivity. Although an ADHD probability index is not calculated for 2- to 5-year-olds, an index score for clinical probability and functional impairment is provided.¹²²

The BRIEF-P measures behaviours associated with executive function in preschool children, aged 2 years to 5 years 11 months. A rating form can be completed by parents, educators, and carers to assess executive function at home and in childcare, kindergarten, or preschool settings. Domains of executive functioning, including working memory, emotional control, inhibition, shifting, and ability to plan/organize, are measured. The BRIEF-P has high internal consistency reliability (0.80–0.95 for the parent version and 0.90–0.97 for the teacher version). Clinical utility of the BRIEF-P to screen for ADHD in young children found that of the five subscales, scores on the 'Inhibit' and 'Working memory' subscales were closely related to ADHD symptoms and correctly identified 86.4% of children in ADHD and typically developing groups.¹⁶⁰ The BADDS and the revised version of the Brown Executive Function/Attention Scales consider problems a child who has ADHD might experience with executive function, including organization; focusing, sustaining, and shifting attention; regulating alertness; regulating emotions; working memory; and monitoring.¹⁵⁸ If scores are elevated, a comprehensive clinical assessment with a clinician is appropriate. Parent and teacher versions are available for children 3 to 7 years of age.

The CBCL/1.5–5 is a broad measure, and part of the Achenbach System of Empirically Based Assessment.¹⁶¹ The CBCL/1.5–5 includes an attention syndrome scale, and a DSM-oriented diagnostic scale for ADHD. Parent and teacher forms, with 99 items, as well as open-ended questions gain qualitative information about a child's problems and strengths. It has been reported to have acceptable specificity (specificity=0.81) and reasonable sensitivity (sensitivity=0.75).¹⁵⁶ The Conners EC assessment tool is used to gather parents' and teachers' observations about behaviour of children aged 2 to 6 years.¹⁵⁹ It assesses behavioural, emotional, social, and developmental issues in young children. The behavioural scales include an inattention/hyperactivity scale. In addition, the Conners EC offers two global index scores that may be useful in screening for ADHD in young children: the Conners EC Global Index: Restless-Impulsive and the Conners EC Global Index: Total. Children with elevated scores on the Conners EC Global Index: Restless-Impulsive may be easily distracted, restless, fidgety, or impulsive. Children with elevated scores on the Conner EC Global Index: Total may be moody and emotional, restless, impulsive, and inattentive. The SDQ¹²⁰ was summarized in screening tools in the social-emotional domain. Scores are provided for five scales: emotional symptoms, conduct problems, inattention/hyperactivity, peer problems, and prosocial behaviour. A total difficulties score is also provided. The hyperactivity/inattention subscale shows good sensitivity (sensitivity=0.8) at the expense of specificity (specificity=0.48).¹⁵⁶ The SDQ has great utility as a brief screening assessment for attentional difficulties. The BASC-3, CBCL, and Conners EC all provide information about attention and behaviour in the context of broader functioning and may also be useful in

screening for ADHD symptoms, while also exploring general social-emotional functioning and development.

DCD

International clinical practice recommendations on assessment, diagnosis, and intervention of DCD¹⁴ recommend children under 5 years are assessed over two time points, at least 3 months apart with a strict criterion of the fifth centile cut-off (severe impairment), and holistic consideration given to the child and their environment.^{14,162} For the children scoring at or below the 16th centile on the MABC-2, monitoring and follow-up are recommended. The MABC-2 is the most cited preschool-age motor assessment for DCD^{14,92,162} and has best predictive validity where 4-year-old children with MABC-2 scoring below the fifth centile are likely to have motor impairment at 8 years.¹⁶² The Developmental Coordination Disorder Questionnaire¹⁶³ is a parent questionnaire which screens for coordination disorders in children aged 5 to 15 years; and the Little Developmental Coordination Disorder Questionnaire screens for motor coordination difficulties in 3- to 4-year-old children.¹⁶³⁻¹⁶⁵

FASD

FASD is a neurodevelopmental disorder associated with prenatal alcohol exposure.^{11,166} The neurodevelopmental impairments affect many aspects of a child's life, including cognitive, motor, language, and social functioning.¹¹ Diagnostic assessment for FASD is best conducted by a multidisciplinary team.¹¹ For a diagnosis of FASD three criteria must be considered: (1) evidence of prenatal alcohol use and other exposures; (2) neurodevelopmental impairment; and (3) specific facial features.^{1,11} Alternative diagnoses that might explain an individual's physical and neurodevelopmental impairments must be excluded. Forms for FASD diagnostic assessment and management planning are provided for clinical use.¹¹ Multiple screening tools are summarized in a 2022 review; however, there were no specific screening tools for FASD recommended for clinical use.¹⁶⁶ General developmental screening tools can be used to screen for potential neurodevelopmental impairments. Assessment should be prioritized according to the child's functional difficulties, age, and local resources.¹¹ Routinely gathering information about prenatal alcohol exposure when taking a developmental history will enable accurate differential diagnosis. Collection and assessment of prenatal alcohol exposure can be undertaken using the Alcohol Use Disorders Identification Test.¹⁶⁷

DISCUSSION

This scoping review identified 67 tools with good discriminative or predictive validity for the screening and assessment of

developmental concerns or disability in children from birth to 5 years 11 months of age. Early accurate identification can enable timely referral to early intervention programmes that can have a positive influence on many outcomes.¹⁶⁸⁻¹⁷⁰ Early accurate identification of children who have developmental concerns or a disability is the key first step to implementing successful early intervention strategies. It is well recognized that, to be effective, screening children for a disability or developmental concern should begin early and be repeated through childhood.^{171,172} Screening involves the use of standardized screening tools to identify children who are at risk and may lead to more in-depth evaluation in which history taking, standardized assessment tools, and clinical observations are used to diagnose a delay or disability. As a child's family know them best, it is vital that families are engaged as partners in the developmental screening and evaluation process.

The best screening tools recommended for use by health professionals and early educators in primary care settings include the following: for multiple developmental domains, the ASQ-3; for autism, the SACS-R, SACS-PR, and ASDetect (smartphone app available for parents/caregivers); for communication, the CSBS; for social-emotional concerns, the ASQ-SE-2 and SDQ. These tools can be used and scored by general practitioners, maternal/child health nurses, health workers, and allied health professionals. These tools may be used in rural and remote areas through telehealth as they are parent questionnaires. The strongest evidence-based screening tool for early detection of CP and other adverse neurodevelopmental outcomes is the GMA.

The screening tool that covers multiple developmental domains and has the best published psychometric evidence is the ASQ-3.²⁸ The ASQ-3 can be used by all health professionals and early educators so it is well-suited to the primary care setting. Parents and health professionals report it is acceptable and understandable.¹⁷³ Parents report the ASQ-3 enables them to 'assess their own child and to work in partnership' with health professionals.¹⁷³ Although training is not mandatory for health professionals, it is advisable for consistency in use and scoring.¹⁷³ Health professionals and early educators need to be aware of the limitations of the ASQ-3 in that it is a screener, and will not identify all children with developmental concerns and may incorrectly identify others. Sensitivity and specificity of the ASQ are moderate but more accurate in older children (>24 months) than in younger ones (<24 months).²³ As a screening tool, the ASQ items do not take into full consideration the quality of performance. For example, a young child may demonstrate asymmetry of movements that the ASQ-3 motor domain may not discern, whereas a comprehensive motor assessment such as the HINE will more fully discern atypical motor performance. Development of clear recommendations for children who score 'close to the cut-off' for two or more domains on the ASQ-3 (e.g. communication, problem-solving, and personal-social) would be beneficial. Intentional, systematic screening using the ASQ-3 at the early developmental

checks would benefit children from disadvantaged populations. The ASQ-3 may be used with parents face-to-face, through telehealth, or online, enabling it to be utilized in urban, regional, and remote primary care settings. Use of a skills-based checklist screener can help parents better understand their child's development.

Three multi-domain developmental assessment tools had strong psychometrics: the Bayley-III,⁴⁰ the BDI-2,⁴¹ and the MSEL.⁴² While the Bayley-III is recognized as a 'criterion standard' in validity studies, it is limited to a ceiling age of 3 years 6 months. The MSEL can be used up to 5 years 8 months, while the BDI-2 extends up to 7 years 11 months. Four multi-domain functional/adaptive assessment tools that discriminate typical versus delayed development of functional performance and adaptive behaviour are recommended: the PEDI,⁵³ PEDI-CAT,⁵⁴ VABS-3,⁵⁵ and ABAS-3.⁵⁶ These four functional assessments are parent-report questionnaires which may be done online, or by interview through telehealth, providing greater flexibility for assessment remotely. The VABS-3 and ABAS-3 are commonly used in diagnosing intellectual disability. None of the multi-domain tools require formal training in test procedure (although it is recommended for the Bayley); however, the tools are restricted to use by qualified health professionals trained in a relevant discipline (e.g. speech pathologist, occupational therapist, physiotherapist, psychologist). Routinely gathering information about prenatal alcohol exposure when taking a developmental history is important for consideration of FASD.

Training and professional support may be required to increase the capability of clinicians and the capacity of the workforce to utilize the tools. Tools recommended for comprehensive assessment of neurodevelopmental outcomes are the HINE and HAI (<2 years) and the NSMDA (>2 years). Motor assessment tools with sound psychometric evidence and strong clinical utility are the PDMS-2 and MABC-2. Gross and fine-motor skill assessment are also contained in the multi-domain assessment tools (Bayley-III, BDI-2, MESEL). Cognition domains are assessed within the multi-domain tools, with the Bayley-III being the most commonly used assessment of cognition for very young children. Cognitive tools with strong psychometrics such as the WPPSI-IV and SB-5 are only available for use by qualified psychologists.

Communication tools are many, and the recommendations for assessment tools in this review focused on tools that provided useful clinical information over multiple communication domains. The Preschool Language Scales, Fifth Edition and CELF-5 are recommended for assessment of language, and the Diagnostic Evaluation of Articulation and Phonology and the Dynamic Evaluation of Motor Speech Skill for speech.

The BITSEA, SDQ, and ASQ-SE-2 can be used for screening of social-emotional concerns. Assessment tools such as the ITSEA, CBCL, and Greenspan Social-Emotional Chart are used for in-depth assessment by suitably qualified health professionals. Tools for ADHD include the BASC-3, BRIEF,

and BADDs. Assessment tools for autism include the ADOS-2, CARS-2, and ADI-R. Clinicians are referred to national guidelines for autism¹³ and ADHD.¹¹ Tools for diagnoses of CP include the GMA and Motor Optimality Score-Revised in early infancy, and the HINE, HAI, and GMFM with increasing age. For children with CP, classification of functional severity is described using the Gross Motor Function Classification System, the Manual Ability Classification System, and the Communication Function Classification System.

Early detection of developmental concerns using validated screening tools that are culturally appropriate for Indigenous communities is important as inappropriate tools may lead to over- or under-recognition of children with developmental concerns.^{17,174,175} Culturally safe methods and tools are more frequently used in First Nations healthcare settings.^{16,174} Four tools were identified that have been culturally validated for the Indigenous children in Australia: the ASQ-TRAK, HATS, PLUM, and the Indigenous Child-Initiated Pretend Play Assessment (with its component of the Indigenous Play Partner Scale; Appendix S4).

This scoping review aimed to provide recommendations on the best evidence-based screening tools and/or developmental and diagnostic assessments to identify children with developmental concerns or a disability across broad international settings. Factors that complicate the ability to make comparisons of tools on the basis of predictive validity or discriminative validity are that statistics of test accuracy are determined by the type of measurement tool, the cut-off points used in a study (e.g. 1 or 2SD below the mean), and the participants in a study (high-risk group vs low-risk group; age). The recommendations given are based on available data from this extensive scoping review which were drawn from systematic review findings. To maintain currency, it is recommended that the scoping review is updated every 5 years.

CONCLUSIONS

This scoping review provides clinicians with recommendations for the best evidence-based tools to assess developmental concerns or to diagnose disability on the basis of predictive or discriminative ability and clinical utility. The best multi-domain screening tool is the ASQ-3. The Bayley-III and BDI-2 are the best multi-domain developmental assessments. Recommended diagnostic assessment tools for CP are the GMA and the HINE. The best diagnostic screening tool for autism is the SACS-R. These tools and assessments have broad utility across multiple settings internationally.

ACKNOWLEDGEMENT

Open access publishing facilitated by The University of Queensland, as part of the Wiley - The University of Queensland agreement via the Council of Australian University Librarians.

FUNDING INFORMATION

Australian Government Department of Health and Aged Care; the Merchant Charitable Foundation via the Children's Hospital Foundation (donation identifier: 10415). RNB is supported by an NHMRC Investigator grant, NHMRC Research Fellowship (1105038 RNB), University of Queensland Research Stimulus Grant (AB), NHMRC Early Career Fellowship (KB, APP1145212), Cerebral Palsy Alliance; PHD02619 (CL).


CONFLICT OF INTEREST STATEMENT

The authors have stated that they had no interests that might be perceived as posing a conflict or bias.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in supporting online information, published studies, and Figshare. <https://doi.org/10.6084/m9.figshare.21708248.v1>.

ORCID

Andrea Burgess  <https://orcid.org/0000-0001-8697-5156>
 Carly Luke  <https://orcid.org/0000-0002-2732-1942>
 Michelle Jackman  <https://orcid.org/0000-0002-5467-1649>
 Jane Wotherspoon  <https://orcid.org/0000-0002-6893-7196>
 Koa Whittingham  <https://orcid.org/0000-0002-5344-9907>
 Katherine Benfer  <https://orcid.org/0000-0002-7252-5063>
 Sarah Reedman  <https://orcid.org/0000-0002-3937-5649>
 Leanne Sakzewski  <https://orcid.org/0000-0001-5395-544X>
 Roslyn N. Boyd  <https://orcid.org/0000-0002-4919-5975>

REFERENCES

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5-TR. 5th edition, text revision. Washington, DC: American Psychiatric Association Publishing; 2022.
2. Australian Institute of Health and Welfare. Australia's children. In: Welfare AIoHa editor. Canberra: Australian Institute of Health and Welfare; 2020.
3. Australian Cerebral Palsy Register Group. Report of the Australian Cerebral Palsy Register, Birth Years 1995–2012. Australian Cerebral Palsy Register, Group; 2018.
4. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, Dan B, Jacobsson B. A Report: the definition and classification of cerebral palsy. *Developmental Medicine & Child Neurology Supplement* 2007; 49: 8–14.
5. Zeidan J, Fombonne E, Scora H, Ibrahim A, Durkin MS, Saxena S, Yusuf A, Shih A, Elabbagh M. Global prevalence of autism: A systematic review update. *Autism Res* 2022; 15: 778–90.
6. Department of Social Services. Australia' Disability strategy 2021–2031: Early childhood Targeted Action Plan. In: Services DoS editor. Canberra; 2021.
7. Peters MDJ, Godfrey C, McInerney P, Khalil H, Larsen P, Marnie C, Pollock D, Tricco AC, Munn Z. Best practice guidance and reporting items for the development of scoping review protocols. *JBIM Evid Synth* 2022; 20: 953–68.

8. Burgess A, RN. B. Tools for early detection of disability and developmental concerns in young children (aged birth-5 years 11 months) in Australia.
9. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, Williamson PR, Terwee CB. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. *Trials* 2016; 17: 449.
10. Australian ADHD Guideline Development Group. Australian Evidence-Based Clinical Practice Guideline for Attention Deficit Hyperactivity Disorder (ADHD). Melbourne, Australia: Australian ADHD Professionals Association; 2022.
11. Bower C, Group. EEobotS. Report to the Australian Government Department of Health: "Australian Guide to the diagnosis of Fetal Alcohol Spectrum Disorder (FASD)" 2016.
12. Verdon S, Mackenzie N, McLeod S, Davidson C, Masso, S., Verdon L, Edwards-Grove C. Assessment of Children as Effective Communicators in Early Childhood Education and Care: Literature Review. Melbourne, VIC: Victorian Curriculum and Assessment Authority; 2018.
13. Whitehouse AJO, Evans K, Eapen V, Wray J. A national guideline for the assessment and diagnosis of autism spectrum disorders in Australia. In: Autism CrCfLw editor. Brisbane; 2018.
14. Blank R, Barnett AL, Cairney J, Green D, Kirby A, Polatajko H, Rosenblum S, Smits-Engelsman B, Sugden D, Wilson P, Vincon S. International clinical practice recommendations on the definition, diagnosis, assessment, intervention, and psychosocial aspects of developmental coordination disorder. *Dev Med Child Neurol* 2019; 61: 242–85.
15. Novak I, Morgan C, Adde L, Blackman J, Boyd RN, Brunstrom-Hernandez J, Cioni G, Damiano D, Darrah J, Eliasson AC, de Vries LS, Einspieler C, Fahey M, Fehlings D, Ferriero DM, Fethers L, Fiori S, Forssberg H, Gordon AM, Greaves S, Guzzetta A, Hadders-Algra M, Harbourne R, Kakooza-Mwesige A, Karlsson P, Krumlinde-Sundholm L, Latal B, Loughran-Fowlds A, Maitre N, McIntyre S, Noritz G, Pennington L, Romeo DM, Shepherd R, Spittle AJ, Thornton M, Valentine J, Walker K, White R, Badawi N. Early, Accurate Diagnosis and Early Intervention in Cerebral Palsy: Advances in Diagnosis and Treatment. *JAMA Pediatr* 2017; 171: 897–907.
16. Ching TYC, Hou S, Seeto M, Harkus S, Ward M, Marnane V, Kong K. The Parents' evaluation of Listening and Understanding Measure (PLUM): Development and normative data on Aboriginal and Torres Strait Islander children below 6 years of age. *Deafness & Education International* 2020; 22: 288–304.
17. Ching TYC, Saetre-Turner M, Harkus S, Martin L, Ward M, Marnane V, Jones C, Collyer E, Khamchuang C, Kong K. The Hearing and Talking Scale (HATS): Development and validation with young Aboriginal and Torres Strait Islander children in urban and remote settings in Australia. *Deafness & Education International* 2020; 22: 305–24.
18. Dender AM, Stagnitti KE. Content and cultural validity in the development of the Indigenous Play Partner Scale. *Aust Occup Ther J* 2017; 64: 283–93.
19. Simpson S, Eadie T, Khoo ST, Titmuss A, Maple-Brown LJ, Thompson R, Wunungmurra A, Jeyaseelan D, Dunham M, D'Aprano A. The ASQ-TRAK: Validating a culturally adapted developmental screening tool for Australian Aboriginal children. *Early Hum Dev* 2021; 163: 105481.
20. Cairney DG, Kazmi A, Delahunty L, Marryat L, Wood R. The predictive value of universal preschool developmental assessment in identifying children with later educational difficulties: A systematic review. *PLoS ONE [Electronic Resource]* 2021; 16.
21. da Silva MA, de Mendonça Filho EJ, Mõnego BG, Bandeira DR. Instruments for multidimensional assessment of child development: A systematic review. *Early Child Development and Care* 2020; 190: 1257–71.
22. Hjorth S, Bromley R, Ystrom E, Lupattelli A, Spigset O, Nordeng H. Use and validity of child neurodevelopment outcome measures in studies on prenatal exposure to psychotropic and analgesic medications - A systematic review. *PLoS ONE [Electronic Resource]* 2019; 14.
23. Muthusamy S, Wagh D, Tan J, Bulsara M, Rao S. Utility of the Ages and Stages Questionnaire to Identify Developmental Delay in Children Aged 12 to 60 Months: A Systematic Review and Meta-analysis. *Journal of American Medical Association Pediatrics* 2022.
24. Pontoppidan M, Niss NK, Pejtersen JH, Julian MM, Væver MS. Parent report measures of infant and toddler social-emotional development: a systematic review. *Family Practice* 2017; 34: 127–37.
25. Schonhaut L, Maturana A, Cepeda O, Seron P. Predictive Validity of Developmental Screening Questionnaires for Identifying Children With Later Cognitive or Educational Difficulties: A Systematic Review. *Front Pediatr* 2021; 9: 698549.
26. van Dokkum NH, Reijneveld SA, de Best JTBW, Hamoen M, Te Wierike SCM, Bos AF, de Kroon MLA. Criterion Validity and Applicability of Motor Screening Instruments in Children Aged 5–6 Years: A Systematic Review. *International Journal of Environmental Research and Public Health* 2022; 19.
27. Velikonja T, Edbrooke-Childs J, Calderon A, Slead M, Brown A, Deighton J. The psychometric properties of the Ages & Stages Questionnaires for ages 2-2.5: a systematic review. *Child: Care, Health & Development* 2017; 43: 1–17.
28. Squires J, Bricker D, Potter L. Revision of a parent-completed development screening tool: ages and stages questionnaires. *J Pediatr Psychol* 1997; 22: 313–28.
29. Squires J, Bricker D. Ages & Stages Questionnaires, Third Edition (ASQ-3): A Parent Completed Child Monitoring System. Baltimore: Paul H. Brookes Publishing Company, Inc.; 2009.
30. Rah SS, Jung M, Lee K, Kang H, Jang S, Park J, Yoon JY, Hong S-B. 4.5 Real-World Accuracy of Developmental Screening Tests for Children: A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry* 2021; 60: S148–S9.
31. D'Aprano A, Institute. MSoHrasftL. Guide to using the ASQ-TRAK: Ages & Stages Questionnaires. Paul H. Brookes Publishing Co.; 2017.
32. D'Aprano A, Silburn S, Johnston V, Robinson G, Oberklaid F, J. S. Adaptation of the Ages and Stages Questionnaire for Remote Aboriginal Australia. *Qualitative Health Research* 2016; 2016: 613–25.
33. Parents' Evaluation of Developmental Status (PEDS). <https://pedstest.com/AboutOurTools/LearnAboutPEDS/IntroductionToPEDS.html> (accessed 25 October 2022)
34. Brothers KB, Glascoe FP, Robertshaw NS, kkk. PEDS: Developmental Milestones—An Accurate Brief Tool for Surveillance and Screening. *Clinical Pediatrics* 2008; 47: 271–9.
35. Alotaibi M, Long T, Kennedy E, Bavishi S. The efficacy of GMFM-88 and GMFM-66 to detect changes in gross motor function in children with cerebral palsy (CP): a literature review. *Disability & Rehabilitation* 2014; 36: 617–27.
36. Johnson S, Bountziouka V, Brocklehurst P, Linsell L, Marlow N, Wolke D, Manktelow BN. Standardisation of the Parent Report of Children's Abilities-Revised (PARCA-R): a norm-referenced assessment of cognitive and language development at age 2 years. *The Lancet Child & Adolescent Health* 2019; 3: 705–12.
37. Johnson S, Marlow N, Wolke D, Davidson L, Marston L, O'Hare A, Peacock J, Schulte J. Validation of a parent report measure of cognitive development in very preterm infants. *Dev Med Child Neurol* 2004; 46: 389–97.
38. Martin AJ, Darlow BA, Salt A, Hague W, Sebastian L, McNeill N, Tarnow-Mordi W. Performance of the Parent Report of Children's Abilities-Revised (PARCA-R) versus the Bayley Scales of Infant Development III. *Arch Dis Child* 2013; 98: 955–8.
39. Kallioinen M, Eadon H, Murphy MS, Baird G, Guideline C. Developmental follow-up of children and young people born preterm: summary of NICE guidance. *BMJ* 2017; 358: j3514.

40. Bayley N. Bayley scales of infant and toddler development. San Antonio, TX: PsychCorp, Pearson; 2006.
41. Newborg J. Battelle Developmental Inventory. Itasca, IL: Riverside; 2005.
42. Mullen EM. Mullen scales of early learning (AGS ed.). Circle Pines, MN: American Guidance Service Inc.; 1995.
43. Griffiths A, Toovey R, Morgan PE, Spittle AJ. Psychometric properties of gross motor assessment tools for children: A systematic review. *BMJ Open* 2018; 8.
44. Spittle AJ, Doyle LW, Boyd RN. A systematic review of the clinimetric properties of neuromotor assessments for preterm infants during the first year of life. *Developmental Medicine & Child Neurology* 2008; 50: 254–66.
45. Cunha ACB, Berkovits MD, Albuquerque KA. Developmental Assessment With Young Children: A Systematic Review of Battelle Studies. *Infants & Young Children: An Interdisciplinary Journal of Early Childhood Intervention* 2018; 31: 69–90.
46. Morgan C, Honan I, Allsop A, Novak I, Badawi N. Psychometric Properties of Assessments of Cognition in Infants With Cerebral Palsy or Motor Impairment: A Systematic Review. *Journal of Pediatric Psychology* 2019; 44: 238–52.
47. Bishop SL, Guthrie W, Coffing M, Lord C. Convergent validity of the Mullen Scales of Early Learning and the differential ability scales in children with autism spectrum disorders. *Am J Intellect Dev Disabil* 2011; 116: 331–43.
48. Farmer C, Golden C, Thurm A. Concurrent validity of the differential ability scales, second edition with the Mullen Scales of Early Learning in young children with and without neurodevelopmental disorders. *Child Neuropsychol* 2016; 22: 556–69.
49. Swineford LB, Guthrie W, Thurm A. Convergent and divergent validity of the Mullen Scales of Early Learning in young children with and without autism spectrum disorder. *Psychol Assess* 2015; 27: 1364–78.
50. Akshoomoff N. Use of the Mullen Scales of Early Learning for the assessment of young children with Autism Spectrum Disorders. *Child Neuropsychol* 2006; 12: 269–77.
51. Jones EJH, Mason L, Begum Ali J, van den Boomen C, Braukmann R, Cauvet E, Demurie E, Hessels RS, Ward EK, Hunnius S, Bolte S, Tomalski P, Kemner C, Warreyn P, Roeyers H, Buitelaar J, Falck-Ytter T, Charman T, Johnson MH, Euser T. Euser: Towards robust measurement of infant neurocognitive predictors of autism across Europe. *Infant Behav Dev* 2019; 57: 101316.
52. Luiz D, Faragher B, Barnard A, Knoesen N, Kotras N, Burns LE, Challis D. Griffiths Mental Development Scales–Extended Revised: 2 to 8 years – analysis manual. Oxford, UK: Hogrefe.; 2006.
53. Haley S, Coster W, Ludlow L, Haltiwanger J, Andrellos P. Pediatric Evaluation of Disability Inventory: development, standardization, and administration manual. Boston: Trustees of Boston University, Health and Disability Research Institute.; 1992.
54. Haley SM, Coster WJ, Dumas HM, Fragala-Pinkham MA, Moed R. Pediatric Evaluation of Disability Inventory Computer Adaptive Test, Development, Standardization and Administration Manual Version 1.3.6. Boston, MA: CRECare LLC; 2012.
55. Sparrow SS, Cicchetti DV, Saulnier CA. Vineland-3: Vineland adaptive behavior scales.. Bloomington, MN: Psychcorp; 2016.
56. Harrison P, Oakland T. Adaptive Behavior Assessment System, 3rd ed. Washington, DC, USA: Western Psychological Services; 2015.
57. Oakland T, Harrison PL. *Adaptive Behavior Assessment System-II*; 2008.
58. Ferre-Fernandez M, Murcia-Gonzalez MA, Espinosa MDB, Rios-Diaz J. Measures of Motor and Functional Skills for Children With Cerebral Palsy: A Systematic Review. *Pediatric Physical Therapy* 2020; 32: 12–25.
59. Mensch SM, Rameckers EAA, Ehteld MA, Evenhuis HM. Instruments for the evaluation of motor abilities for children with severe multiple disabilities: A systematic review of the literature. *Research in Developmental Disabilities* 2015; 47: 185–98.
60. Prechtl HFR, Einspieler C, Cioni G, Bos AF, Ferrari F, Sontheimer D. An early marker for neurological deficits after perinatal brain lesions. *The Lancet* 1997; 349: 1361–3.
61. Haataja L, Mercuri E, Regev R, Cowan F, Rutherford M, Dubowitz V, Dubowitz L. Optimality score for the neurologic examination of the infant at 12 and 18 months of age. *The Journal of pediatrics* 1999; 135: 153–61.
62. Krumlinde-Sundholm L, Ek L, Sicola E, Sjostrand L, Guzzetta A, Sgandurra G, Cioni G, Eliasson AC. Development of the Hand Assessment for Infants: evidence of internal scale validity. *Dev Med Child Neurol* 2017; 59: 1276–83.
63. Burns YR. *N.S.M.D.A. Physiotherapy Assessment for Infants and Young Children*. Brisbane, Queensland: Copyright Publishing Company; 2014.
64. Bosanquet M, Copeland L, Ware R, Boyd R. A systematic review of tests to predict cerebral palsy in young children. *Developmental Medicine & Child Neurology* 2013; 55: 418–26.
65. Burger M, Louw QA. The predictive validity of general movements - A systematic review. *European Journal of Paediatric Neurology* 2009; 13: 408–20.
66. Craciunoiu O, Holsti L. A Systematic Review of the Predictive Validity of Neurobehavioral Assessments During the Preterm Period. *Physical & Occupational Therapy in Pediatrics* 2017; 37: 292–307.
67. Darsaklis V, Snider LM, Majnemer A, Mazer B. Predictive validity of Prechtl's Method on the Qualitative Assessment of General Movements: a systematic review of the evidence. *Developmental Medicine & Child Neurology* 2011; 53: 896–906.
68. Seesahai J, Luther M, Church PT, Maddalena P, Asztalos E, Rotter T, Banihani R. The assessment of general movements in term and late-preterm infants diagnosed with neonatal encephalopathy, as a predictive tool of cerebral palsy by 2 years of age—a scoping review. *Systematic Reviews* 2021; 10.
69. Caesar R, Colditz PB, Cioni G, Boyd RN. Clinical tools used in young infants born very preterm to predict motor and cognitive delay (not cerebral palsy): a systematic review. *Developmental Medicine and Child Neurology* 2021; 63: 387–95.
70. Spittle AJ, Olsen J, Kwong A, Doyle LW, Marschik PB, Einspieler C, Cheong J. The Baby Moves prospective cohort study protocol: using a smartphone application with the General Movements Assessment to predict neurodevelopmental outcomes at age 2 years for extremely preterm or extremely low birthweight infants. *BMJ Open* 2016; 6: e013446.
71. Einspieler C, Bos AF, Kriebler-Tomantschger M, Alvarado E, Barbosa VM, Bertoncelli N, Burger M, Chorna O, Del Secco S, DeRegnier RA, Huning B, Ko J, Lucaccioni L, Maeda T, Marchi V, Martin E, Morgan C, Mutlu A, Nokolova A, Pansy J, Peyton C, Pokorny FB, Prinsloo LR, Ricci E, Saini L, Scheuchnegger A, Silva CRD, Soloveichick M, Spittle AJ, Toldo M, Utsch F, van Zyl J, Vinals C, Wang J, Yang H, Yardimci-Lokmanoglu BN, Cioni G, Ferrari F, Guzzetta A, Marschik PB. Cerebral Palsy: Early Markers of Clinical Phenotype and Functional Outcome. *J Clin Med* 2019; 8.
72. Crowle C, Jackman M, Morgan C. The General Movements Motor Optimality Score in High-Risk Infants: A Systematic Scoping Review. *Pediatric Physical Therapy* 2023; 35: 2–25.
73. Kwong AK, Boyd RN, Chatfield MD, Ware RS, Colditz PB, George JM. Early Motor Repertoire of Very Preterm Infants and Relationships with 2-Year Neurodevelopment. *J Clin Med* 2022; 11.
74. Benfer KA, Novak I, Morgan C, Whittingham K, Khan NZ, Ware RS, Bell KL, Bandaranayake S, Salt A, Ghosh AK, Bhattacharya A, Samanta S, Moola G, Bose D, Tripathi S, Boyd RN. Community-based parent-delivered early detection and intervention programme for infants at high risk of cerebral palsy in a low-resource country (Learning through Everyday Activities with Parents (LEAP-CP): protocol for a randomised controlled trial. *BMJ Open* 2018; 8: e021186.
75. Amiel-Tison C. Update of the amiel-tison neurologic assessment for the term neonate or at 40 weeks corrected age. *Pediatric Neurology* 2002; 27: 196–212.

76. Goo M, Tucker K, Johnston LM. Muscle tone assessments for children aged 0 to 12 years: a systematic review. *Developmental Medicine & Child Neurology* 2018; 60: 660–71.
77. Noble Y, Boyd R. Neonatal assessments for the preterm infant up to 4 months corrected age: a systematic review. *Developmental Medicine & Child Neurology* 2012; 54: 129–39.
78. Romeo DM, Ricci D, Brogna C, Mercuri E. Use of the Hammersmith Infant Neurological Examination in infants with cerebral palsy: a critical review of the literature. *Dev Med Child Neurol* 2016; 58: 240–5.
79. Romeo DM, Cowan FM, Haataja L, Ricci D, Pede E, Gallini F, Cota F, Brogna C, Vento G, Romeo MG, Mercuri E. Hammersmith Infant Neurological Examination for infants born preterm: predicting outcomes other than cerebral palsy. *Dev Med Child Neurol* 2021; 63: 939–46.
80. Hay K, Nelin M, Carey H, Chorna O, Moore-Clingenpeel MAMASM, Maitre N. Hammersmith Infant Neurological Examination Asymmetry Score Distinguishes Hemiplegic Cerebral Palsy From Typical Development. *Pediatric Neurology* 2018; 87: 70–4.
81. Pietruszewski L, Moore-Clingenpeel M, Moellering GCJ, Lewandowski D, Batterson N, Maitre NL. Predictive value of the test of infant motor performance and the Hammersmith infant neurological examination for cerebral palsy in infants. *Early Human Development* 2022; 174.
82. Wagenaar N, Verhage CH, de Vries LS, van Gasselt BPL, Koopman C, Leemans A, Groenendaal F, Benders M, van der Aa NE. Early prediction of unilateral cerebral palsy in infants at risk: MRI versus the hand assessment for infants. *Pediatr Res* 2020; 87: 932–9.
83. Ullenhag A, Ek L, Eliasson AC, Krumlinde-Sundholm L. Interrater and test-retest reliability of the Hand Assessment for Infants. *Dev Med Child Neurol* 2021; 63: 1456–61.
84. Ek L, Eliasson AC, Sicola E, Sjostrand L, Guzzetta A, Sgandurra G, Cioni G, Krumlinde-Sundholm L. Hand Assessment for Infants: normative reference values. *Dev Med Child Neurol* 2019; 61: 1087–92.
85. Piper MC, Darrah J. *Motor Assessment of the Developing Infant*. Philadelphia: W B Saunders; 1994.
86. Folio MR, Fewel RR. *Peabody developmental motor scales: Examiners' manual*. Austin: Pro-Ed; 2000.
87. Henderson SE, Sugden DA, Barnett AL. *Movement assessment battery for children-2: Movement ABC-2: Examiner's manual*. Pearson; 2007.
88. Ulrich DA. *Test of gross motor development-2*. Austin: Pro-Ed; 2000.
89. Bruininks R, Bruininks B. *Bruininks-Oseretsky Test of Motor Proficiency-2nd Edition (BOT-2)*. Circle Pines: MN: AGS Publishing; 2005.
90. Albuquerque PLd, Lemos A, Guerra MQdF, Eickmann SH. Accuracy of the Alberta Infant Motor Scale (AIMS) to detect developmental delay of gross motor skills in preterm infants: A systematic review. *Developmental neurorhabilitation* 2015; 18: 15–21.
91. De Roubaix A, Van de Velde D, Roeyers H, Van Waelvelde H. Standardized motor assessments before the age of five predicting school-aged motor outcome including DCD: A systematic review. *European Journal of Paediatric Neurology* 2021; 30: 29–57.
92. Bieber E, Smits-Engelsman BC, Sgandurra G, Cioni G, Feys H, Guzzetta A, Klingels K. Manual function outcome measures in children with developmental coordination disorder (DCD): Systematic review. *Res Dev Disabil* 2016; 55: 114–31.
93. Nicola K, Waugh J, Charles E, Russell T. The feasibility and concurrent validity of performing the Movement Assessment Battery for Children - 2nd Edition via telerehabilitation technology. *Res Dev Disabil* 2018; 77: 40–8.
94. Gill K, Osioviich A, Synnes A, J AA, Grunau RE, Miller SP, Zwicker JG. Concurrent Validity of the Bayley-III and the Peabody Developmental Motor Scales-2 at 18 Months. *Phys Occup Ther Pediatr* 2019; 39: 514–24.
95. Soler N, Cordier R, Perkes IE, Dale RC, Bray P. Proxy-reported sensory measures for children and adolescents with neurodevelopmental disorders: A systematic review. *Developmental Medicine & Child Neurology* 2022; 30: 30.
96. Eeles AL, Spittle AJ, Anderson PJ, Brown N, Lee KJ, Boyd RN, Doyle LW. Assessments of sensory processing in infants: a systematic review. *Developmental Medicine & Child Neurology* 2013; 55: 314–26.
97. Pfeiffer B, Coster W, Tucker C, Piller A, 2018; OTMH. Development and Content Validity of the Participation and Sensory Environment Questionnaire. 34 2018; 2: 105–21.
98. Dunn W. *Sensory Profile 2: User's Manual*. TX: Psych Corporation; 2014.
99. Wong HS, Santhakumaran S, Cowan FM, Modi N. Developmental Assessments in Preterm Children: A Meta-analysis. *Pediatrics* 2016; 138: 1–12.
100. DiStefano CD, S.C. *Journal of psychoeducational assessment.. Investigating the Theoretical Structure of the Stanford-Binet-Fifth Edition..* 2006 2006; 24: 123–36.
101. Korkman M, Kirk U, Kemp S. NEPSY-II Review. *Journal of Psychoeducational assessment* 2007; 28: 175–82.
102. Marshall S, McGoe KE, Moschos S. Test Review. *Journal of Psychoeducational Assessment* 2011; 29: 89–93.
103. Renaud F, Beliveau MJ, Akzam-Ouellette MA, Jauvin K, Labelle F. Comparison of the Wechsler Preschool and Primary Scale of Intelligence-Third Edition and the Leiter-R Intellectual Assessments for Clinic-Referred Children. *J Psychoeduc Assess* 2022; 40: 825–38.
104. Yin Foo R, Guppy M, Johnston LM. Intelligence assessments for children with cerebral palsy: a systematic review. *Developmental Medicine & Child Neurology* 2013; 55: 911–8.
105. Wechsler D. *The Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV) Technical and Interpretive Manual Supplement: Special Group Validity Studies With Other Measures*. USA: Pearson; 2012.
106. Roid G, Miller L. *Leiter international performance scale-revised: Examiner's manual*. Wood Dale, IL: Stoelting; 1997.
107. Ruffini C, Tarchi C, Morini M, Giuliano G, Pecini C. Tele-assessment of cognitive functions in children: A systematic review. *Child Neuropsychology* 2022; 28: 709–45.
108. Wetherby AM, Prizant BM. *Communication and symbolic behavior scales: Developmental profile..* Baltimore, MD:: Paul H Brookes Publishing; 2002.
109. Sansavini A, Favilla ME, Guasti MT, Marini A, Millepiedi S, Di Martino MV, Vecchi S, Battajon N, Bertolo L, Capirci O, Carretti B, Colatei MP, Frioni C, Marotta L, Massa S, Michelazzo L, Pecini C, Piazzalunga S, Pieretti M, Rinaldi P, Salvadorini R, Termine C, Zuccarini M, D'amico S, De Cagno AG, Levorato MC, Rossetto T, Lorusso ML. Developmental language disorder: Early predictors, age for the diagnosis, and diagnostic tools. A scoping review. *Brain Sciences* 2021; 11.
110. Salgado-Cacho JM, Moreno-Jimenez MDP, de Diego-Otero Y. Detection of Early Warning Signs in Autism Spectrum Disorders: A Systematic Review. *Children (Basel)* 2021; 8.
111. Strand EA, McCauley RJ, Weigand SD, Stoeckel RE, Baas BS. A motor speech assessment for children with severe speech disorders: Reliability and validity evidence. *Journal of Speech, Language, and Hearing Research*, 2013; 56: 505–20.
112. Gubiani MB, Pagliarin KC, Keske-Soares M. Tools for the assessment of childhood apraxia of speech. *CoDAS (São Paulo)* 2015; 27: 610–5.
113. Dodd B, Corporation. P. *Diagnostic evaluation of articulation and phonology: DEAP*. London, England: Psychological Corporation Ltd.; 2002.
114. Zimmerman IL, Steiner VG, Pond RE, Karas G, Brian MJ, Boland J, Marshall E. *PLS-5 preschool language scales. 5th ed. (Australian and New Zealand adaptation)*. Sydney, N.S.W.: Pearson Clinical and Talent Assessment; 2012.

115. Wiig EH, Semel EM, Secord W. CELF-Preschool: clinical evaluation of language fundamentals--Preschool (2nd ed.). San Antonio: Psychological Corp., Harcourt Brace Jovanovich.; 2004.
116. Briggs-Gowan MJ, Carter AS, Irwin JR, Wachtel K, Cicchetti DV. The brief infant-toddler social and emotional assessment: screening for social-emotional problems and delays in competence. *J Pediatr Psychol* 2004; 29: 143–55.
117. Carter AS, Briggs-Gowan MJ, Jones SM, Little TD. The Infant-Toddler Social and Emotional Assessment (ITSEA): Factor structure, reliability, and validity. *Journal of abnormal child psychology* 2003; 31: 495–514.
118. Squires JK, Bricker D, Twombly E. The ASQ-SE User's Guide: For the Ages & Stages Questionnaires, Social-Emotional. A Parent-Completed, Child-Monitoring System for Social-Emotional Behaviors. Baltimore, MD: Paul H. Brookes Pub. Co.; 2002.
119. Goodman R. The Strengths and Difficulties Questionnaire. *Journal of child Psychology and Psychiatry* 1997; 38: 581–6.
120. Goodman R. Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry* 2001; 40: 1337–45.
121. Hanratty J, Livingstone N, Robalino S, Terwee CB, Glod M, Oono IP, Rodgers J, Macdonald G, McConachie H. Systematic review of the measurement properties of tools used to measure behaviour problems in young children with autism. *PLoS ONE [Electronic Resource]* 2015; 10.
122. Reynolds CR, RW. K. Behavior Assessment System for Children.: Bloomington: Pearson; 2015.
123. Achenbach TM, Rescorla LA. Manual for the ASEBA preschool forms and profiles. Burlington, VT: Department of Psychiatry, University of Vermont; 2000.
124. Bayley N. Greenspan Social-Emotional Growth Chart in Bayley Scales of Infant Development and Toddler Development: Technical Manual.. San Antonio, TX: The PsychCorp; 2006.
125. Tede Z, Ophir Cohen M, Riskin A, Tirosh E. The reliability and validity of the Greenspan Social Emotional Growth Chart (GSEGC) in Israeli children with developmental delay and autism-A pilot study. *Res Dev Disabil* 2016; 55: 226–34.
126. Dender A, Stagnitti K. Development of the Indigenous Child-Initiated Pretend Play Assessment: Selection of play materials and administration. *Aust Occup Ther J* 2011; 58: 34–42.
127. Voress JK, Maddox T. Developmental assessment of young children (2nd ed.). Austin, TX: PRO-ED.; 2013.
128. Palisano RJ, Rosenbaum P, Bartlett D, Livingston MH. Content validity of the expanded and revised Gross Motor Function Classification System. *Dev Med Child Neurol* 2008; 50: 744–50.
129. Palisano RJ, Rosenbaum P, Walter SD, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol* 1997; 39: 214–23.
130. Eliasson AC, Krumlinde-Sundholm L, Rosblad B, Beckung E, Arner M, Ohrvall A-M, Rosenbaum P. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Developmental Medicine & Child Neurology* 2006; 48: 549–54.
131. Hidecker MJC, Paneth N, Rosenbaum PL, Kent RD, Lillie J, Eulenberg JB, et al. Developing and validating the Communication Function Classification System for individuals with cerebral palsy. *Developmental medicine and child neurology* 2011; 53: 704–10.
132. Russell DJ, Hart HMP. Gross motor function measure (GMFM-66 & GMFM-88) user's manual. London: Mac Keith Press; 2013.
133. Rosenbaum PL, Walter SD, Hanna SE, Palisano RJ, Russell DJ, Raina P, Wood E, Bartlett DJ, Galuppi BE. Prognosis for Gross Motor Function in Cerebral Palsy: Creation of Motor Development Curves. *JAMA* 2002; 288: 1357–63.
134. Burgess A, Reedman S, Chatfield MD, Ware RS, Sakzewski L, Boyd RN. Development of gross motor capacity and mobility performance in children with cerebral palsy: a longitudinal study. *Dev Med Child Neurol* 2022; 64: 578–85.
135. Barbaro J, Sadka N, Gilbert M, Beattie E, Li X, Ridgway L, Lawson LP, Dissanayake C. Diagnostic Accuracy of the Social Attention and Communication Surveillance-Revised With Preschool Tool for Early Autism Detection in Very Young Children. *JAMA Netw Open* 2022; 5: e2146415.
136. Robins DL, Casagrande BS, Barton M, Chen C, 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.
137. Barbaro J, Yaari M. Study protocol for an evaluation of ASDetect - a Mobile application for the early detection of autism. *BMC Pediatr* 2020; 20: 21.
138. Hirota T, So R, Kim YS, Leventhal B, Epstein RA. A systematic review of screening tools in non-young children and adults for autism spectrum disorder. *Research in Developmental Disabilities* 2018; 80: 1–12.
139. Mawle E, Griffiths P. Screening for autism in pre-school children in primary care: systematic review of English Language tools. *International Journal of Nursing Studies* 2006; 43: 623–36.
140. Mozolic-Staunton B, Donnelly M, Yoxall J, Barbaro J. Early detection for better outcomes: Universal developmental surveillance for autism across health and early childhood education settings. *Research in Autism Spectrum Disorders* 2020; 71.
141. Petrocchi S, Levante A, Lecciso F. Systematic review of level 1 and level 2 screening tools for autism spectrum disorders in toddlers. *Brain Sciences* 2020; 10.
142. Sánchez-García AB, Galindo-Villardón P, Nieto-Librero AB, Martín-Rodero H, Robins DL. Toddler Screening for Autism Spectrum Disorder: A Meta-Analysis of Diagnostic Accuracy. *Journal of Autism & Developmental Disorders* 2019; 49: 1837–52.
143. Yuen T, Penner M, Carter MT, Szatmari P, Ungar WJ. Assessing the accuracy of the Modified Checklist for Autism in Toddlers: a systematic review and meta-analysis. *Developmental Medicine & Child Neurology* 2018; 60: 1093–100.
144. Aishworiya R, Van Kim M, Stewart S, Hagerman R, Feldman HM. Meta-analysis of the Modified Checklist for Autism in Toddlers, Revised/Follow-up for Screening.. *Pediatrics* 2023; 151.
145. Geng X, Kang X, Wong PCM. Autism spectrum disorder risk prediction: A systematic review of behavioral and neural investigations. *Progress in Molecular Biology and Translational Science* 2020; 173: 91–137.
146. Breidbord J, Croudace TJ. Reliability Generalization for Childhood Autism Rating Scale. *Journal of Autism & Developmental Disorders* 2013; 43: 2855–65.
147. Lebersfeld JB, Swanson M, Clesi CD, O'Kelley SE. Systematic Review and Meta-Analysis of the Clinical Utility of the ADOS-2 and the ADI-R in Diagnosing Autism Spectrum Disorders in Children. *Journal of Autism & Developmental Disorders* 2021; 51: 4101–14.
148. Moon SJ, Hwang JS, Shin AL, Kim JY, Bae SM, Sheehy-Knight J, Kim JW, Sheehy-Knight J. Accuracy of the Childhood Autism Rating Scale: a systematic review and meta-analysis. *Developmental Medicine & Child Neurology* 2019; 61: 1030–8.
149. Randall M, Egberts KJ, Samtani A, Scholten RJPM, Hooft L, Livingstone N, Sterling-Levis K, Woolfenden S, Williams K. Diagnostic tests for autism spectrum disorder (ASD) in preschool children. *Cochrane Database of Systematic Reviews* 2018; 2018.
150. Lord C. Autism Diagnostic Observation Schedule: ADOS-2: manual. Los Angeles (CA): Western Psychological Services; 2012.
151. Schopler E, Reichler RJ, BR. R. CARS (The childhood autism rating scale).. Los Angeles: Western Psychological Services; 1988.
152. Kim SH, Lord C. New autism diagnostic interview-revised algorithms for toddlers and young preschoolers from 12 to 47 months of age. *J Autism Dev Disord* 2012; 42: 82–93.
153. Bryson SE, Zwaigenbaum L, McDermott C, Rombough V, Brian J. The Autism Observation Scale for Infants: scale development and reliability data. *J Autism Dev Disord* 2008; 38: 731–8.
154. Chang LY, Wang MY, Tsai PS. Diagnostic accuracy of Rating Scales for attention-deficit/hyperactivity disorder: A meta-analysis. *Pediatrics* 2016; 137.

155. Hourii AK, Miller FG. A Systematic Review of Universal Screeners Used to Evaluate Social-Emotional and Behavioral Aspects of Kindergarten Readiness. *Early Education and Development* 2019; 31: 653–75.
156. Mulraney M, Arrondo G, Musullulu H, Iturmendi-Sabater I, Cortese S, Westwood SJ, Donno F, Banaschewski T, Simonoff E, Zuddas A, Döpfner M, Hinshaw SP, Coghill D. Systematic Review and Meta-analysis: Screening Tools for Attention-Deficit/Hyperactivity Disorder in Children and Adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry* 2022; 61: 982–96.
157. Gioia GA, Isquith PK, Guy SC, Kenworthy L. Behavior rating inventory of executive function (BRIEF): Professional manual. Lutz, Florida: Psychological Assessment Resources; 2000.
158. Brown TE. Attention Deficit Disorder: The Unfocused Mind in Children and Adults. New Have: Yale University Press.; 2005.
159. Conners CK. The Conners 3rd Edition (Conners 3). Toronto, Ontario, Canada.: Multi-Health Systems.; 2008.
160. Skogan AH, Zeiner P, Egeland J, Urnes A, Reichborn-Kiennerud T, Aase H. Parent ratings of executive function in young preschool children with symptoms of attention-deficit/hyperactivity disorder.. *Behav Brain Funct* 2015; 11: 16.
161. al Ae. Reliability and Validity of CBCL. 2019.
162. Lee EJ, Zwicker JG. Early identification of children with/at risk of developmental coordination disorder: a scoping review. *Dev Med Child Neurol* 2021; 63: 649–58.
163. Wilson BN, Crawford SG. The Developmental Coordination Disorder Questionnaire 2007 - Manual. Calgary, Alberta, Canada; 2012.
164. Gabbard C, Tamplain P. The Strengths and Limitations of DCD-Related Screening Questionnaires. *Current Developmental Disorders Reports* 2021; 8: 1–5.
165. Wilson BN, Creighton D, Crawford SG, Heath JA, Semple L, Tan B, Hansen S. Psychometric Properties of the Canadian Little Developmental Coordination Disorder Questionnaire for Preschool Children. *Phys Occup Ther Pediatr* 2015; 35: 116–31.
166. Lim YH, Watkins RE, Jones H, Kippin NR, Finlay-Jones A. Fetal alcohol spectrum disorders screening tools: A systematic review. *Research in Developmental Disabilities* 2022; 122: 104168.
167. Bush K, Kivlahan DR, McDonell, M. B., Fihn SD, Bradley KA. The AUDIT Alcohol Consumption Questions (AUDIT-C): An Effective Brief Screening Test for Problem Drinking. *Archives of Internal Medicine* 1998; 158: 1789–95.
168. Reynolds AJ, Temple JA, Ou S, al. e. Effects of a School-Based, Early Childhood Intervention on Adult Health and Well-being: A 19-Year Follow-up of Low-Income Families.. *Arch Pediatr Adolesc Med*. 2007; 161: 730–9.
169. Spittle A, Orton J, Anderson PJ, Boyd R, Doyle LW. Early developmental intervention programmes provided post hospital discharge to prevent motor and cognitive impairment in preterm infants. *Cochrane Database Syst Rev* 2015; 2015: CD005495.
170. Whitehouse AJO, Varcin KJ, Pillar S, Billingham W, Alvares GA, Barbaro J, Bent CA, Blenkley D, Boutrus M, Chee A, Chetcuti L, Clark A, Davidson E, Dimov S, Dissanayake C, Doyle J, Grant M, Green CC, Harrap M, Iacono T, Matys L, Maybery M, Pope DF, Renton M, Rowbottom C, Sadka N, Segal L, Slonims V, Smith J, Taylor C, Wakeling S, Wan MW, Wray J, Cooper MN, Green J, Hudry K. Effect of Preemptive Intervention on Developmental Outcomes Among Infants Showing Early Signs of Autism: A Randomized Clinical Trial of Outcomes to Diagnosis. *JAMA Pediatr* 2021; 175: e213298.
171. Cairney J, Clark HJ, Nair K. Parental Concerns, Developmental Temperature Taking, and the Necessary Conditions for Developmental Surveillance and Screening. *Current Developmental Disorders Reports* 2016; 3: 174–9.
172. Lipkin PH, M.M. M, Disabilities. CoCW. Promoting Optimal Development: Identifying Infants and Young Children With Developmental Disorders Through Developmental Surveillance and Screening. *Pediatrics* 2020; 145: e20193449.
173. Kendall S, Nash A, Braun A, Bastug G, Rougeaux E, Bedford H. Acceptability and understanding of the Ages & Stages Questionnaires(R), Third Edition, as part of the Healthy Child Programme 2-year health and development review in England: Parent and professional perspectives. *Child Care Health Dev* 2019; 45: 251–6.
174. Kendall E, Barnett L. Principles for the development of Aboriginal health interventions: culturally appropriate methods through systemic empathy. *Ethnicity & Health* 2014; 20: 437–52.
175. Simpson S, D'Aprano A, Tayler C, Toon Khoo S, Highfold R. Validation of a culturally adapted developmental screening tool for Australian Aboriginal children: Early findings and next steps. *Early Hum Dev* 2016; 103: 91–5.
176. Roid GH, Pomplun M. The Stanford-Binet Intelligence Scales, *Fifth Edition*. Itasca, IL: Riverside Pub; 2012.
177. Elliott CD. The Differential Ability Scales-Second Edition. San Antonio, TX: Harcourt Assessment; 2007.

SUPPORTING INFORMATION

The following additional material may be found online:

Appendix S1: Search Strategies for systematic reviews and grey literature.

Appendix S2: Excel Table - Summary of Psychometric Properties and Clinical Utility.

Appendix S3: List of tools with abbreviations.

Appendix S4: Culturally appropriate tools for Indigenous Australian children.

Appendix S5: Systematic reviews included in review.

Figure S1: Flowchart.

How to cite this article: Burgess A, Luke C, Jackman M, Wotherspoon J, Whittingham K, Benfer K, et al. Clinical utility and psychometric properties of tools for early detection of developmental concerns and disability in young children: A scoping review. *Dev Med Child Neurol*. 2025;67:286–306. <https://doi.org/10.1111/dmcn.16076>