BMC Primary Care

RESEARCH

Open Access

Brief multidimensional screening tools for young children's mental health and development for administration by primary care providers: a scoping review

Ashley D Radomski[1,2*], Christine Polihronis[1,2,3], Paula Cloutier[2,4], Kayla Beaudin[1,2] and Mario Cappelli[1,2,3,5,6]

Abstract

Introduction Many young children with mental health or developmental concerns go unidentified and untreated, underscoring the need for effective screening. Primary care offers unique opportunities for screening, yet rates are low. Primary care providers often cite time constraints, insufficient training, and costs as barriers to screening. Tools designed for administration by primary care providers hold promise for facilitating personalized assessments and improving communication, collaboration, and follow-up between providers and families. To improve early identification and intervention, it is crucial to understand provider-administered screening tools that align with their practical constraints.

Aim To identify and describe the characteristics, limitations, and improvement areas of brief, multidimensional screening for young children's mental health and development, designed for administration by primary care providers.

Methods We conducted a scoping review according to published guidelines. We searched seven electronic databases and used hand-searching strategies. We sought English-language publications on screening tools (English or French) assessing at least one mental health and one development domain in children up to 6 years, designed for administration by primary care providers in under 20 minutes. Two reviewers assessed the articles' eligibility and then extracted, charted, and summarized relevant data.

Results Three screening tools from six articles were included. The tools were primarily administered by physicians or nurses to children aged 2 weeks to 4 years and required minimal training to use. The tools varied in their domains (4–6), items (10–110), psychometric properties, and scoring methods, but all included indicators of delayed or at-risk children. Article limitations included study recency and lack of data related to implementation and patient outcomes.

Conclusions This review described the development and evaluation of multidimensional screening tools for young children's mental health and development designed for administration by primary care providers. It found that few tools have been published. This review identifies several knowledge gaps and emphasizes the need for research on

*Correspondence:
Ashley D Radomski
aradomski@cheo.on.ca

Full list of author information is available at the end of the article

the implementation, cost-effectiveness, and comparative performance of screening tools, as well as the development and evaluation of new screening tools that suit providers' needs. Future studies should investigate how these screening tools can improve identification, clinical care, and health outcomes.

Keywords Screening tool, Primary care, Health care providers, Review, Mental health, Child development, Child, Infant screening, Assessment, Pediatrics

Introduction

Major gaps exist between the proportion of young children experiencing mental health and/or developmental concerns, those who are correctly identified, and those who receive timely and appropriate care [1–4]. Mental health and development in infants and young children includes various aspects of physical and mental well-being, such as linguistic, cognitive, and socio-emotional functioning, all of which depend on safe and supportive relationships, environments, and nutrition to foster their growth [5]. Considerable evidence demonstrates that early identification and intervention can significantly improve children's long-term outcomes [6, 7], with various assessment and treatment options available [8–11]. Clinical guidelines worldwide recognize the importance of screening and the unique opportunity primary care provides to initiate this process [12–20].

Primary care includes the provision of first-line, accessible healthcare services for the majority of personal health needs, along with continuous, person-focused care over time, and coordinated entry into the health system for needs and problems requiring specialized care [21–24]. Primary care is often the first point of contact for the healthcare of young children [25], particularly those with mental health and/or developmental concerns [4, 26]. Primary care providers, including family physicians, pediatricians, and nurses, are well-positioned to cultivate trusting relationships with families, conduct ongoing monitoring of children, guide caregivers on aspects of child health, and coordinate connections with relevant resources and services [4, 14, 25, 27, 28]. Primary care providers may be encouraged to screen young children for mental health and/or developmental concerns based on a recommended schedule, such as well-child visits [12, 15, 19, 29], during any visit when symptoms, problems, or risk factors are identified, or when concerns from a parent, early childhood professional, or health care provider are raised [13, 30]. Key aspects of young children's mental health and development—cognition, language, motor, social, and emotion regulation abilities—are often assessed together due to their complex interplay [5, 31–33]. However, there is no standard definition of mental health and development for this age group, nor agreement on which domains are essential to assess.

Screening is a systematic assessment approach that commonly involves the use of an evidence-based tool, such as an interview tool, behavioural assessment, questionnaire, or rating scale. Screening tools serve multiple purposes– they help identify a child's signs and symptoms, assess their severity, and determine the need for further assessment or intervention by comparing results to established norms or threshold scores [34]. Screening practices can be categorized into two approaches: 'universal' screening, where as many children as possible within a given population are assessed, and 'targeted' screening, which focuses on at-risk children or high-risk groups [35]. On the spectrum of assessment methods, screening occupies a middle ground. Screening is more focused and sensitive than milestone tracking, developmental monitoring, or surveillance, but less comprehensive and specific than diagnostic testing [4, 36]. The use of standardized and validated screening tools improves the identification of children [4, 37–40], compared to relying on parent concerns, observation, or clinical judgment alone [4, 40–43]. Screening young children before diagnostic testing serves two important purposes. It helps allocate diagnostic resources efficiently and identifies children who may need clinical support even if they do not meet formal diagnostic criteria [14].

While the benefits of screening are widely acknowledged [15, 19, 44, 45], the use of standardized screening tools falls considerably below recommended levels, ranging from 23 to 48% [46–49]. Primary care providers report barriers such as a lack of time, validated tools, training, and remuneration to conduct screening [40, 46–56]. Furthermore, when young children are identified, the rates of referral or follow-up care are also low [39, 45, 57, 58], resulting in few children receiving the timely and appropriate care they need. Low referral and follow-up rates may be related, in part, to inadequate resources, funding, referral or service options [59], but also to whether screening tools are incorporated into clinical visits and used to inform the next steps in care [44, 57].

Although no 'gold standard' screening tool has been established [38], previous literature has identified several tools for assessing the mental health or development of young children [4, 38, 57, 60–64]. Many widely adopted tools, including the Ages & Stages Questionnaire [65] and Parents' Evaluation of Developmental Status– Revised [66], are designed for parents to complete independently without direct involvement from a primary care provider [4, 9, 43, 67]. While parent-completed tools can save providers time during visits with families [68, 69], they may

be insufficient when parents or providers require immediate or additional information and follow-up.

Provider-administered screening tools offer an alternative to parent-completed assessments. With these tools, the primary care provider conducts the screening during the clinical visit, gathering information directly from the parent (or caregiver) and/or the child. Provider-administered tools can enhance information gathering by providing opportunities for conversation, direct observation, interactions between the child, caregiver, and provider, and addressing concerns in real time. These tools may better align with caregiver assessment preferences and elicit caregiver disclosure or mitigate caregiver language or literacy challenges as the provider can rephrase questions or provide helpful prompts or examples. In addition, provider-administered screening tools may clarify perceptions and understanding of mental health and developmental behaviors and concerns and help the provider consider important contextual information relevant to the child's assessment and care [56, 70, 71]. This approach also allows providers to introduce concepts related to child mental health and development, along with the purpose of screening, discuss findings with caregivers, provide caregiver support, develop follow-up plans, and assess the family's readiness or preferences for services, all within a single visit [14, 72]. Consequently, care may be improved through increased collaboration, communication, and shared decision-making, leading to more effective, person-centred care [73, 74].

Improving the screening and referral rates of young children in primary care requires a thorough examination of available screening tools for primary care providers. These tools should align with providers' time, training, and cost constraints [45, 75–79]. To address this need, we conducted a scoping review of published, brief, multidimensional screening tools for the mental health and development of young children, designed for administration by primary care providers. Our review aimed to identify existing tools, describe their characteristics and applications, highlight potential limitations, and suggest areas for improvement. A greater understanding of available screening tools and their characteristics can be used to strengthen screening practices in primary care and facilitate earlier identification and treatment of children with mental health and developmental concerns.

Methods

Study design

Our approach to this scoping review was informed by commonly used methodological and reporting guidelines, including Arksey and O'Malley's framework for conducting scoping reviews [80, 81], JBI methodology for scoping reviews [82, 83] and Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) extension for scoping reviews [84, 85]. The protocol was published with Open Science Framework (OSF) Registries [86].

Search strategy and information sources

This scoping review aimed to include published, peer-reviewed articles on brief, multidimensional screening tools administered by primary care providers to identify both mental health and developmental concerns in young children. Importantly, tools used for developmental milestone tracking, monitoring, or surveillance, like the Rourke Baby Record [87, 88] and Looksee Checklist (formerly the Nipissing District Developmental Screening) [89], or comprehensive assessments, such as the Bayley Scales of Infant Development [90], were out of scope for this review.

Electronic search strategies were developed by a Research Librarian (KF) and peer-reviewed (MCD) using the PRESS guideline [91] (Additional file 1). We searched 7 electronic databases: MEDLINE(R) ALL (OvidSP), Embase (OvidSP), Cochrane Controlled Register of Trials (OvidSP), CINAHL (EBSCOhost), APA PsycINFO (OvidSP), ERIC (OvidSP) and PsycTESTS (APA PsycNET). Starting from their inception, each database was searched for the concepts of "screening tool", "child development/mental health" and "primary care provider" using a combination of subject headings and keywords. We included only English-language articles. Non-peer-reviewed research articles, such as commentaries, editorials, and conference proceedings, were removed. To validate the initial search strategy, 100 abstracts were retrieved using Ovid MEDLINE and reviewed by the authorship team in August 2019. A Research Librarian used the articles identified as included (true positives) and excluded (true negatives) to refine and update the search strategy. We conducted our initial comprehensive electronic search using all selected databases at 3 time points: (1) September 17, 2019, (2) January 15, 2021, and (3) July 12, 2022. We conducted hand searching with the reference lists of relevant reviews and included articles. The results from the various searches were imported into Covidence software [92] for the screening and selection process. We used Zotero [93] as the reference management software and Microsoft Office 365 Word for data charting and tracking the methods, procedures, and decisions of this review.

Literature screening and selection

The screening and selection process was performed in two stages by a combination of two independent reviewers (ADR, AR, or HV). The first stage involved reviewing each article's title and abstract for eligibility. If the eligibility of an abstract was unclear, article screening was conducted. The second stage involved reviewing

the article for inclusion or exclusion. Reasons for exclusion were noted in Covidence and are presented in the PRISMA flow diagram (Fig. 1). Disagreements in inclusion or exclusion decisions were resolved via discussion between two reviewers and other review team members. Deduplication was initially performed by a Research Librarian, who employed the search strategy and ran an automatic process in Covidence.

Eligibility criteria

Seven eligibility criteria were used to determine whether articles were included or excluded (Table 1). We included primary and secondary published, peer-reviewed articles that described and/or evaluated brief, multidimensional mental health and developmental screening tools for young children up to 6 years of age, designed to be administered by primary care providers.

Contacting corresponding authors

We contacted the corresponding authors of included or potentially eligible articles to identify and resolve missing information. If no response was received after at least 2 contact attempts, team consensus was used to make relevant decisions about an article's eligibility and/or the data to be charted. Out of the eight authors contacted, three responded.

Inter-reviewer reliability

We used the most common quantitative method of calculating inter-reviewer reliability, Cohen's kappa, to demonstrate the magnitude of agreement between two independent reviewers on a random subset of articles during the two-stage screening process [100, 101]. Reviewer 1 (ADR) remained constant throughout screening, whereas Reviewer 2 (AR) and Reviewer 3 (HV) each

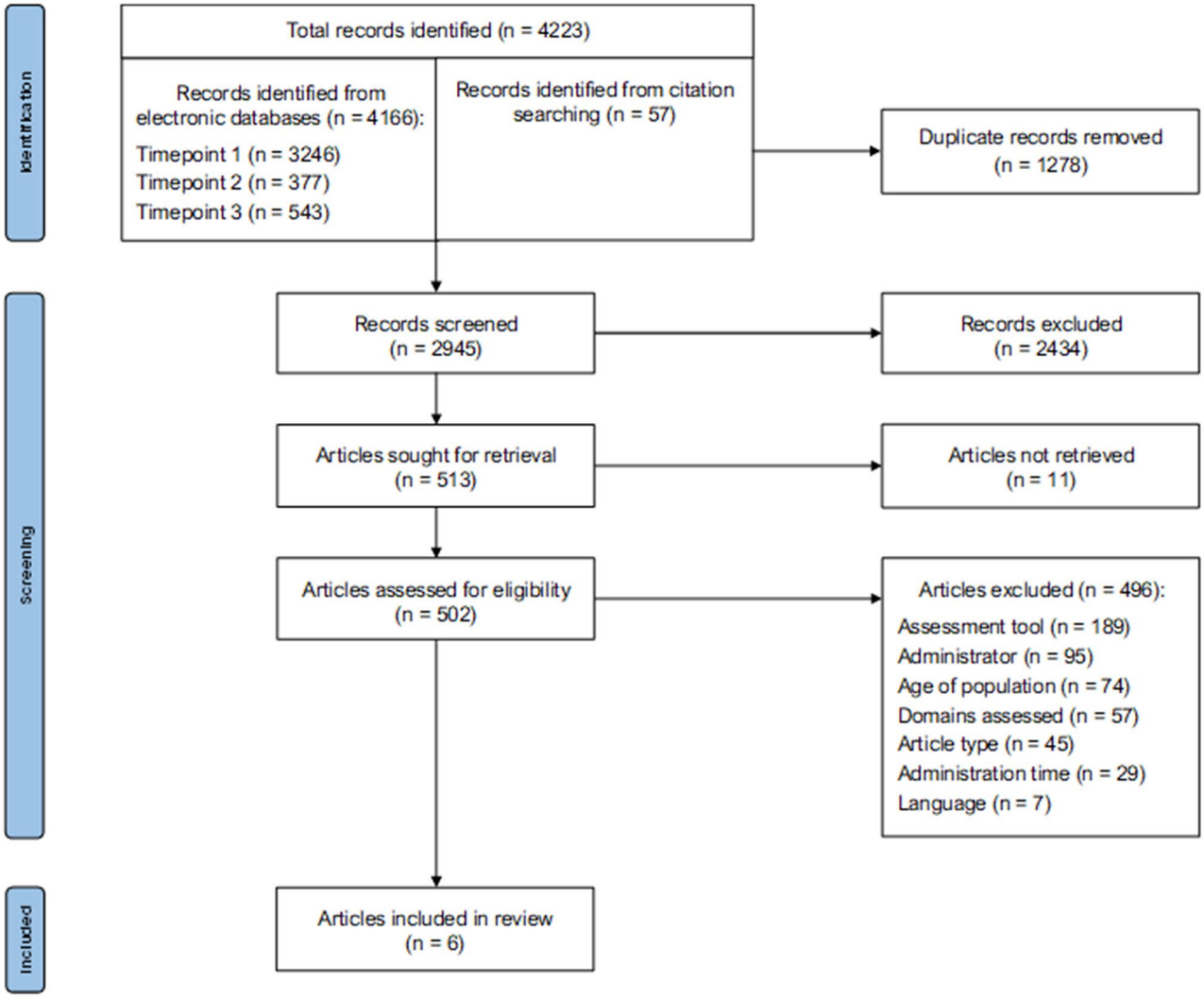

Fig. 1 PRISMA flow diagram of our literature search and selection process
Timepoint 1: September 17, 2019, Timepoint 2: January 15, 2021, Tmepoint 3: July 12, 2022

Table 1 Inclusion and exclusion criteria used to determine the eligibility of articles

Criteria	Inclusion	Exclusion
1. Article type	• Peer-reviewed publication • Primary or secondary research article	• Unpublished and/or non-peer-reviewed article (e.g., grey literature) • Not based on a primary or secondary research study (e.g., review)
2. Language	• Article was in English • Screening tool was available in English or French	• Article was not in English • Screening tool was not available in English or French
3. Screening tool	• Screening tool	• Surveillance, monitoring, or diagnostic assessment tool
4. Age of population	• Assessed children up to 6 years of age	• Assessed children older than 6 years of age
5. Domains assessed	• Assessed at least one domain of mental health and one domain of development at all ages • Mental health domains could include social, emotional, interpersonal, and self-regulation abilities [11, 94] • Developmental domains could include cognition, language, literacy, motor skills, and sensory abilities [95–97]	• Assessed either mental health or development (not both) • Assessed only one domain of functioning (e.g., language) • Assessed a specific health condition (e.g., autism) • Assessed academic aptitude or school readiness
6. Administration time	• Administration time was 'brief' and within the limits of a typical clinical visit (≤ 20 minutes) [75, 98,99]	• Administration time was lengthy and extended beyond the time of a typical clinical visit (> 20 minutes) • Administration of the screening tool could not be isolated from a larger battery of tests or comprehensive assessment
7. Administrator	• Screening tool was designed to be administered or completed by a primary care provider, although other healthcare providers or practitioners may also administer the tool • Primary care providers could include a family physician, pediatrician, nurse practitioner [25, 54] • Researchers or other healthcare practitioners may administer the tool during its development or evaluation	• Screening tool was designed to be administered or completed by a professional not considered to be a primary care provider • Screening tool solely based on parent or caregiver-report

screened a different subset of articles. Inter-reviewer reliability was 'strong' for both titles and abstracts (Reviewers 1 and 2: $n = 100$, $\kappa = 0.89$; Reviewers 1 and 3: $n = 50$, $\kappa = 0.92$) and full-text articles screened (Reviewers 1 and 2: $n = 50$, $\kappa = 0.91$; Reviewers 1 and 3: $n = 20$, $\kappa = 0.90$).

Data extraction and charting

Reviewers extracted and charted data using comprehensive data tables in Microsoft Office 365 Word. The templates for the tables were generated by the review team based on recommendations in the literature [83, 102]. We incorporated a training stage where the reviewers developed and tested their approach to data extraction and charting (for 3–5 articles). Then, reviewers divided and charted the remaining data independently. Charted data was initially reviewed by one reviewer (ADR) before being reviewed by the larger team.

Data items

Extracted data was charted according to 4 main categories: **(1) Article characteristics**: authors, year, country of study, study design and aim(s), full citation; **(2) Participant or population characteristics**: sample size, age, sex, specific health/condition factors; **(3) Screening tool characteristics**: tool name, an overview of the tool and its development, administration details, number of items and domains assessed, administration time, score(s) and/or outcome(s) generated, recommendations provided, cost and training requirements, language or cultural adaptations and availability of the tool; and **(4) Psychometric properties of the screening tool**: comparator instrument(s), measures and outcomes of the screening tool's validity, reliability, sensitivity and specificity, other outcomes relevant to the testing or implementation of the tool.

Critical appraisal of information sources

Following JBI and method guidelines, an assessment of the methodological quality of included articles was not performed for this scoping review [103].

Data synthesis

Descriptive summaries were generated by the review team based on the extracted data. Text and summary tables were developed to identify and describe each screening tool, compare key characteristics between tools, and highlight similarities, differences, potential limitations, and areas for improvement among included screening tools.

Results

Literature search and selection of tools

The PRISMA flow diagram illustrates our search and selection process [84] (Fig. 1). We identified 4223 records from a combination of electronic database and hand-searching strategies. After duplicates were removed, 2945

records underwent title and abstract screening. Following the initial exclusion process, 513 articles were sought for retrieval and 502 articles were obtained and assessed for eligibility. After screening, only six articles describing three screening tools were included in the review.

Description of included articles and screening tools

Table 2 provides a brief overview of the main characteristics of three screening tools, as reported in six included articles spanning the years 1966 to 2010. The screening tools, along with the number of associated articles and their respective countries of origin, are as follows: (1) Brigance Infant and Toddler Screens: Two articles from the United States [72, 104]; (2) Developmental Screening Inventory: One article from the United States [105] and two articles from Nigeria [106, 107]; and (3) Woodside System Screening Technique: One article from Australia [108]. Further details about the screening tools and articles are included in Additional file 2.

Development and validation of screening tools

The Brigance Infant and Toddler Screens were developed in response to consumer requests that the original Brigance Screens [109–111] extend their assessment to younger children. A screen for infants, 2 weeks to 11 months of age, and a screen for toddlers, 12 to 24 months of age, were developed and initially validated in the United States in both English and Spanish [104]. A multidisciplinary panel selected specific items from the larger Brigance Inventory of Early Development [112], pilot-tested these screens on volunteers, and further refined them before publication [104]. Although previous Brigance Screens were designed to be administered by parents or teachers [109, 110, 113], the Brigance Infant and Toddlers Screens were designed to be administered by healthcare providers, such as pediatricians and psychologists. A version of the screens was also developed for independent parent reports.

The Developmental Screening Inventory (DSI) was originally developed in English for children aged 4 weeks to 18 months [105]. It was first validated in the United States for children between 16 and 65 weeks [105], followed by Britain and other locations [114] (article irretrievable). Decades later, the DSI was validated in Nigeria for children between 8 weeks to 30 months of age [106]. The DSI was developed by selecting items from each domain of the Gesell Developmental Schedules [115], and evolved from a similar tool, the 40-week questionnaire, also created by the original authors of the DSI [116]. The DSI was designed for administration by a healthcare provider, who would select age-appropriate questions from all available items. This design allowed the tool to be used for screening children at multiple age intervals [105].

The Woodside System Screening Technique (WSST) was developed in English for children between birth and 5 years of age in Scotland [117] (article irretrievable). The WSST was designed for administration by child health nurses in the preventive care of children up to 5 years old, ensuring easy integration into their clinical routines [108]. The tool was later modified, namely the gross motor domain, and validated with children between 6 weeks to 4 years of age in Australia [108].

Although each screening tool was presented in at least one of the included articles, none of the tools or their technical manuals have been located for formal access or purchase online.

Number of items

The total number of items reported for the three screening tools ranged from 10 to 110; however, each tool provided some form of age-specific items—particular items relevant to the child's developmental stage or age range so that not all possible items are administered. A 2002 article on the Brigance Infant and Toddler Screens reported that each screen included 81 to 85 items [104] whereas a 2010 article reported 90 to 110 items for each year of age [72]. The DSI provided between 10 and 14 items per age group [105], with age groups defined at 4-week intervals. The WSST had a total of 70 items and outlined a stepwise approach for administering them. Age-appropriate reference points were assigned to the items, guiding administrators on where to start and facilitating efficient assessment of whether a child's development status was above, below, or matched to their age. A child's developmental level was determined based on successfully completing a minimum of 2 consecutive items in each domain of the tool [108].

Domains assessed

The screening tools assessed between 4 and 6 domains. Table 3 shows that while there was some variability in the mental health and developmental domains assessed by these tools, all of them had four domains in common: social skills, fine motor, gross motor, and language.

Participant characteristics

Participant samples included both healthy and at-risk children, with nearly equal proportions of males and females. Samples of healthy children were described in one article on the Brigance Infant and Toddler Screens (total $N = 408$; females $n = 196$, 48%) [104] and two articles on the DSI using the same sample (total $N = 128$; females $n = 61$, 47.7%) [106, 107]. Another article on the Brigance Infant and Toddlers Screens included a sample of healthy children with psychosocial or developmental risk factors (total $N = 382$; females $n = 183$, 48%) [72], and one article on the DSI reported on a similar sample (total $N = 58$; sex

Table 2 Overview of the characteristics of included screening tools

Screening tool, article references, country	Age of children[a]	Domains assessed	Number of items	Administration time (minutes)[b]	Scores and indicators	Guidelines and recommendations on care	Sensitivity and/or specificity	Available languages	Cost	Training methods	Availability of the tool
Brigance Infant and Toddler Screens Glascoe (2002) [104]; Glascoe and Leew (2010) [72] USA	Infant screen: 2 weeks to 12 months Toddler screen: 12 to 24 months	1. Fine motor 2. Gross motor 3. Receptive language 4. Expressive language 5. Self-help[c] 6. Socio-emotional	81–110[d]	10–15	Cut-off and standardized domain and total scores	Yes	Sensitivity: Infant screen: 77%; Toddler screen: 76% Specificity: Infant screen: 86%; Toddler screen: 85%	English, Spanish	N/A	Training manual	In article(s) Not formally supported nor available for purchase online[e]
Developmental Screening Inventory (DSI) Knobloch, Pasamanick, and Sherard (1966) [105]; Aina and Morakinyo (2001) [106]; Aina and Morakinyo (2005) [107] USA, Nigeria	4 weeks to 30 months	1. Fine motor 2. Gross motor 3. Language 4. Adaptive 5. Personal-social	10–14[d]	~ 20	'At risk' indicators and standardized domain scores	No	Over-screened: 6.25–25% Under-screened: 0–5%	English	N/A	Instruction manual	In article(s) Not formally supported nor available for purchase online
Woodside System Screening Technique (WSST) Eu (1986) [108] Australia	6 weeks to 4 years	1. Vision and fine motor 2. Gross motor 3. Hearing and language 4. Social	70	~ 15–20	"Normal', 'doubtful', and 'abnormal' indicators for each domain	Yes	Sensitivity: 81% Specificity: 94%	English	N/A	95% inter-rater agreement on 10 consecutive administrations	In article Not formally supported nor available for purchase online

[a]Age range of children used in validation studies of the screening tool

[b]Reported or estimated administration time

[c]Self-help skills were assessed in children older than 12 months of age

[d]Total available items, including age-specific ones. Not every item may have been administered to every child

[e]Updated versions of Brigance tools are formally available online; however, these tools are designed for administration by educators and the provider-administered Brigance Infant and Toddler Screens are no longer supported

Table 3 Domains of mental health and development assessed by each screening tool

Screening Tool	Mental Health Domains			Developmental Domains			
	Social	Emotion	Other	Language	Gross Motor	Fine Motor	Other
Brigance Infant and Toddler Screens	✓	✓	Self-help	✓	✓	✓	
Developmental Screening Inventory	✓		Adaptive	✓	✓	✓	
Woodside System Screening Technique	✓			✓	✓	✓	Hearing; Vision

not reported) [105]. The WSST article included a small sample of children aged 22 and 52 months from a spastic centre (total $N = 17$; females $n = 8$, 47.1%) and a considerably larger comparison group of children aged 6 weeks to 4 years from varied socioeconomic backgrounds (total $N = 444$; females $n = 220$, 49.5%) [108]. Participant characteristics are further detailed in Additional file 2.

Administration details

All screening tools were administered using paper-based forms. Although the tools were designed for administration by primary care providers, these providers were not designated as exclusive administrators. For instance, the Brigance Infant and Toddler Screens, designed for professionals [104], were administered by primary care providers, such as pediatricians, pediatric nurses, and nurse practitioners, as well as early childhood teachers and psychologists [72, 104]. Similarly, the DSI, designed for healthcare professionals [106, 107], was administered by researchers [106, 107], medical staff and students [105]. Lastly, the WSST was administered by child health nurses [108]. Further administration details are described in Additional file 2.

The screening tools were predominantly administered in primary care settings, including clinics, hospitals, and home visits, with some administrations occurring in alternative settings, such as daycare centres (Brigance Infant and Toddler Screens [72, 104]; DSI [106, 107]) and schools (DSI [106, 107]).

Screening tools required 10 to 20 minutes for administration, close to the approximate time of a typical clinical visit [75, 98, 99]. The Brigance Infant and Toddler Screens reported a completion time of 10–15 minutes [104]. Administration of the DSI was estimated to take 20 minutes or less [106], considering the number of test items and its comparison to the Bayley Scales of Infant Development [118]. The WSST was described as less time-consuming than the Denver Developmental Screening Test [108], which typically requires 15–20 minutes for administration [119]. The articles did not provide information on whether screening tools were designed to be incorporated as part of routine and/or problem (specific) visits, nor the time required to score or interpret the screening tools.

Scoring, outcomes, and interpretation

Articles provided information on each screening tool's scoring procedures and outcomes, including how results could be used to identify children with delayed behaviors or who are at risk of a major concern or disorder (see Additional file 2 for more details). For the Brigance Infant and Toddler Screens, total, cut-off, and standardized scores could be generated for all items and items within each domain. Standardized scores that fell below the 16th percentile cut-off (or one standard deviation below the mean) indicated a possible delay and need for diagnostic assessment. An "At-Risk Guideline" accompanied this tool to aid clinicians in diagnostic assessment decisions [72, 104].

The DSI required calculating a child's maturity age and developmental quotient (DQ) for each of the five domains. The maturity age was the maximal age at which the child could carry out a developmental task. The DQ was the ratio of maturity age to chronological age multiplied by 100. A DQ score less than 75 was considered 'questionable' or 'abnormal'; a score above 75 was considered 'adaptive' [105–107]. No specific guidance was provided on how the DQs or other scores could guide clinical care.

The WSST charted the developmental status of each child according to their age and the developmental level of the activities they successfully completed. The tool featured two developmental step patterns (trajectories) for each domain: (1) an expected (age-appropriate), and (2) a threshold (potentially delayed) pattern of development. Children were considered 'normal' if a score reached or exceeded the expected development pattern, 'doubtful' if a score fell between expected and threshold development, and 'abnormal' if a score fell below the threshold development pattern. Referral for a diagnostic assessment was recommended if: (a) 1 or more abnormal scores were recorded, (b) 2 or more doubtful scores were recorded, or (c) a child scored normal initially but doubtful or abnormal at a later assessment [108].

Cost, training, and availability of screening tools

The articles did not provide information on the cost of the screening tools; however, they did include details on the training or competency required for each tool's administration. It was reported in articles on the Brigance Infant and Toddler Screens [104] and DSI [105] that administrators could use the instruction manual to

become familiar with the tool. While professionals could administer the DSI without specific training in developmental diagnoses, administrators often received a brief introduction to the tool and common developmental issues [105–107]. Training for the WSST was considered complete when a trainee achieved an inter-rater agreement score of 95% with a trainer across 10 consecutive administrations [108].

A copy of each screening tool was included in at least one published article (Brigance Infant and Toddler Screens [104]; DSI [105, 106]; WSST [108]). Each tool was available in English; however, in two articles on the Brigance Infant and Toddler Screens, a Spanish version of the tool was also used [72, 104].

Psychometric properties

Four articles described the psychometric properties of the three screening tools, including their validity, reliability, sensitivity, and/or specificity. For additional details, see Additional file 3.

Validity

All screening tools included a measure of concurrent validity through comparison with at least one existing, well-validated, comprehensive assessment tool. Greater correlation coefficients indicate greater agreement between the screening tool and its comparator measure; thus, the concurrent validity is inherently influenced by which comparator measure is used. The most common comparator was the Bayley Scales of Infant Development [118]. In addition to the Bayley Scales of Infant Development, the Brigance Infant and Toddlers Screens were validated against the Infant Behaviour Record, Receptive-Expressive Emergent Language Test, Preschool Language Scale, Alberta Infant Motor Scale, Rosetti Infant Toddler Language Scale, and Vineland Adaptive Behavior Scale [104]. Correlations between the Brigance Screens and related comparators ranged from 0.24 to 0.89 on the infant screen and 0.16 to 0.70 on the toddler screen. Regarding the infant screen, weak correlations (0.24 to 0.41) were reported for comparator measures of daily living, gross motor, and social-emotional skills. Moderate correlations (0.50 to 0.67) were reported for expressive language and cognition. Finally, strong correlations (0.78 to 0.89) were reported for fine motor, global motor development, and receptive language. Regarding the toddler screen, weak correlations (0.16 to 0.40) were reported for comparator measures of daily living skills, social-emotional skills, gross motor, and cognition. Moderate correlations (0.50 to 0.63) were reported for expressive language and global motor development. Finally, strong correlations (0.70) were reported for receptive language and fine motor. These results suggest that comparator measures of receptive language and fine motor had the highest agreement with both the infant and toddler screens. Similarly, both infant and toddler screens had the lowest agreement with comparator measures of daily living, social-emotional, and gross motor skills. Overall, compared with the toddler screen, the infant screen had greater agreement with the comparator measures [104].

The DSI, administered by both student and expert staff examiners, was validated against the Gessell Developmental Examination [105], administered solely by expert staff. Approximately 70% agreement was achieved between both tools; however, in 37 out of 58 cases, when students' administration of the DSI was considered 'adequate' (observation and scoring were systematic and complete), there was 85% agreement. When students administered the DSI, significant correlations were reported between domains of the Gessell Developmental Examination and three DQs from the DSI: adaptive ($r = 0.72$), gross motor ($r = 0.81$), and fine motor ($r = 0.63$) ($p < 0.01$). When the DSI was retroactively completed by staff 2 years following the Gesell Developmental Examination, significant correlations were reported between domains of the Gessell Developmental Examination and five DQs from the DSI: adaptive ($r = 0.97$), gross motor ($r = 0.98$), fine motor ($r = 0.94$), language ($r = 0.97$), and personal-social ($r = 0.97$) ($p < 0.001$). The DSI was also validated against the Bayley Scales of Infant Development [106], showing significant correlations between the Bayley's Psychomotor Developmental Index and the DSI's gross motor, fine motor, language, and personal-social DQs (ranging from r = -0.24 to 0.39). However, there were no significant correlations between the Bayley's Mental Developmental Index and any DQ from the DSI.

The WSST was validated against a diagnostic assessment instrument, the Griffiths Mental Developmental Scales [120, 121]. The Griffiths' Scales were administered to a subset of 124 children by two psychologists, one pediatrician, and the primary author, within 2 weeks of the nurses' administration of the WSST. Significant concurrent validity was reported between the total scores of each tool ($r = 0.73$, $p < 0.05$) and between each WSST domain and the corresponding Griffiths' scale: social ($r = 0.46$), vision and fine motor ($r = 0.65$), hearing and speech ($r = 0.69$), and gross motor ($r = 0.76$) ($p < 0.001$).

Construct validity, which assesses whether a screening tool accurately measures its intended concepts, was only reported in one article for the Brigance Infant and Toddler Screens [104]. A factor analysis revealed a two-factor solution for verbal and non-verbal constructs, accounting for 63% of the variance across items.

Reliability

Interrater reliability measures the agreement between different administrators using the same tool. Interrater reliability of 80% or a Cohen's kappa value > 0.80 is

typically recommended for high agreement [122]. Test-retest reliability measures the agreement between administrations of the same tool at different time points. Ideal test-retest scores are >0.4 for intraclass correlation coefficients, > 0.3 for Pearson correlation coefficients, and >0.4 for Cohen's kappa [123]. Interrater and test-retest reliability were provided for two screening tools. The Brigance Infant and Toddler Screens were readministered by the same examiner and a different examiner within 1 week of the original administration. Test-retest and interrater reliability coefficients ranged from 0.98 to 0.99 for both screens, suggesting high test-rest reliability and interrater agreement.

The WSST showed 95% interrater reliability between two administrators on a subset of 50 children [108] and 95.5% test-retest reliability by the same administrator one week later on a subset of 30 children [108], suggesting high interrater agreement and test-retest reliability. Interrater and test-retest reliability were not reported for the DSI.

Internal consistency, or the degree to which different items within the same screening tool are related, was reported for two tools. Coefficients of ≥0.7 are considered reliable, while coefficients between 0.6 and 0.7 are marginally reliable [123]. The Brigance Infant and Toddler Screens showed Guttman's lambda ranging from 0.94 to 0.97, suggesting high internal consistency, across the infant and toddler screens [104]. The DSI reported a Cronbach's alpha of 0.64, suggesting marginal internal consistency, and provided correlations between raw and DQ scores for all DSI domains [106]. Articles on the WSST did not report internal consistency.

Sensitivity and specificity

The sensitivity and specificity of each screening tool are presented in Table 2. Sensitivity indicates the accuracy of a tool in identifying 'at risk' children who meet diagnostic criteria ("true positive"). Low sensitivity can lead to the under-identification of children at risk. Specificity indicates the accuracy of a tool in identifying children 'not at risk' who do not meet diagnostic criteria ("true negative"). Low specificity can lead to the overidentification of children not at risk. It is typically recommended to maintain sensitivity and specificity thresholds above 70% [9, 124].

For Brigance Infant and Toddler Screens, cut-off scores for developmental delays and diagnoses were generated using receiver operating characteristic analyses and then used to calculate the sensitivity and specificity of the tool. Sensitivity was 77% for the infant screen and 76% for the toddler screen [104]. Specificity was 86% for the infant screen and 85% for the toddler screen.

The WSST was compared to the Griffiths Mental Developmental Scales [120] and demonstrated a sensitivity of 81% and a specificity of 94%. The WSST generated both over-referrals and under-referrals 4% of the time (5/124 children) [108].

Articles on the DSI did not report measures of sensitivity or specificity but did report rates of false negatives (under screened) and false positives (over screened), ranging from 0 to 5% and 6 to 25%, respectively [105].

Other measures

The usability of the DSI screening tool was reported in one article, which found that approximately 64% of medical students, without prior training on the tool, were able to use it adequately [105]. None of the articles on the Brigance Infant and Toddler Screens or WSST described usability. No outcomes were reported regarding either the implementation aspects of these screening tools (such as their feasibility, acceptability, uptake, and cost-effectiveness) or their clinical utility (including their impact on care delivery and child health outcomes).

Discussion

Principal findings

This scoping review aimed to identify and describe published, brief, multidimensional screening tools administered by primary care providers for assessing young children's mental health and development. Six articles that described three screening tools were included. The screening tools commonly assessed 4 to 6 domains of child mental health and development in children aged 2 weeks to 4 years. Primary care physicians and nurses were the main administrators of these tools, with assessments taking between 10 and 20 minutes to complete. Studies of the Brigance Infant and Toddler Screens and the DSI revealed additional flexibility in the administration of their tool, demonstrating successful administration by diverse professionals across diverse settings, even within a single study.

Psychometric properties varied across screening tools. The Brigance Infant and Toddler Screens and the WSST reported high interrater agreement and test-retest reliability. The Brigance Infant and Toddler Screens also reported high internal consistency, while the DSI reported marginal internal consistency. Finally, both the Brigance Infant and Toddler Screens and the WSST reported acceptable sensitivity and specificity thresholds. The Brigance Infant and Toddler Screens had varied concurrent validity outcomes that depended on the comparator instrument used; however, the strongest concurrent validity was reported for receptive language and fine motor domains, while the lowest concurrent validity was reported for daily living, social-emotional, and gross motor. The DSI also reported varied concurrent validity outcomes across domains when compared to the Bayley Scales of Infant Development; greater concurrent

validity was reported when the DSI was compared to the Gessell Developmental Examination. Finally, the WSST reported strong concurrent validity when compared to the Griffiths Mental Developmental Scales.

The evaluation of these screening tools reveals several key considerations. First, while all tools were validated against at least one 'gold standard' measure, their psychometric properties are inherently linked to the chosen comparator instruments. Second, the validity of each tool must be interpreted within the context of its validation population. Notably, none of the studies directly compared outcomes between typically developing and at-risk children, making it difficult to determine whether these tools are better suited for universal population screening or targeted assessment of high-risk groups. All three tools used standardized scoring systems—including cut-off points and risk indicators—to identify children needing further evaluation. Yet only the Brigance Infant and Toddler Screens took the additional step of providing concrete guidance for using these scores to inform diagnostic and referral decisions [104]. Training for administrators was minimal. While a copy of each tool was embedded in at least one article, no articles reported the costs of obtaining or incorporating the tools into clinical practice.

The Brigance Infant and Toddler Screens were translated into Spanish [72], and the DSI was used outside of a high-income country (Nigeria) [106, 107]. One DSI article assessed the usability of the tool among a group of untrained medical students [105]. The articles lacked critical information about real-world implementation, including the tools' acceptability and feasibility in clinical settings. No data were reported on provider or caregiver experiences, nor practical aspects of implementation in primary care or other clinical environments.

Limitations of included screening tools

This review highlights several gaps and limitations in the literature on brief, multidimensional screening tools for the mental health and development of young children, designed to be administered by primary care providers. The first limitation is their temporal distribution or recency. Articles were published between 1966 and 2010, with a lack of recent studies. Additionally, none of the screening tools appear to be currently available for formal access or purchase online. This raises further questions about the acceptance and applicability of these tools in current research and practice, their adaptation to modern advancements in the field (e.g., clinical contexts, guidelines, provider roles), and the reasons for the lack of updated reports, which could indicate important strengths or weaknesses about these tools.

The second limitation concerns the lack of information on implementing the screening tools, including their feasibility, acceptability, and cost-effectiveness. Key practical aspects were not reported, such as scoring time and interpretation, associated costs (e.g., purchasing, billing, remuneration), accessibility (e.g., public domain availability, usage consent), and implementation requirements (e.g., degree of clinical integration, incorporation into a single clinical visit). These factors are crucial, as time and costs are among the most cited barriers to using standardized screening tools [47, 125]. This information is critical for helping providers assess the feasibility and resource implications of adopting a particular screening tool.

Third, the articles provided limited information on how screening results or cut-off scores were applied in clinical care. While the Brigance Infant and Toddler Screens [104] and WSST [108] provided guidelines or recommendations for using outcomes to make diagnostic assessment decisions, no articles reported on providers' adherence to these guidelines or their clinical decisions. The impact of these decisions on children's future care and health outcomes was also not described. Understanding the tools' role in both identification and intervention is crucial for assessing their true utility. This evidence gap significantly hampers primary care providers' ability to make informed implementation decisions, potentially resulting in inefficient resource use, missed early intervention opportunities, or unnecessary assessments that may not benefit young children's mental health or development.

Future research

This scoping review identifies several key areas for improvement in provider-administered screening tools for young children's mental health and development. First, there is a need for future studies to comprehensively evaluate the time required for using screening tools, encompassing administration, scoring, interpretation, and their integration into clinical care. Detailed information on these aspects will provide a clearer understanding of the practical implications of employing screening tools in diverse clinical contexts. Such insights are crucial for primary care providers to efficiently incorporate these tools into their practice and align them with existing clinical workflows.

Second, implementing these screening tools in real-world settings remains largely unexplored. Future research should focus on how the tools are integrated into current practices, including whether they are used how they were designed, highlighting the differences, advantages, and disadvantages across different administrators, settings, and patients. In addition, future research should assess how provider-administered screening tools may be used in tandem with tools completed by other raters, such as a parent, to capture multiple perspectives

and limit biases. Investigating the practical challenges and benefits associated with the application of screening tools will offer valuable insights that can inform best practices and improve the overall efficacy of these tools in everyday healthcare settings.

Third, there is a critical need for economic evaluations or cost-effectiveness analyses of screening tools in clinical practice. Future studies should document the costs of using these tools, including staff, resources, and materials, and undertake cost-effectiveness analyses. Understanding how the costs align with broader health services and the subsequent outcomes for children will provide a holistic perspective on the economic implications of these screening strategies [126]. This information is essential for policymakers and healthcare providers to make informed decisions about allocating resources for screening young children.

A final area of important research should involve comparative studies between different screening tools, particularly parent-completed versus provider-administered tools. Such studies should evaluate administration methods, child and provider characteristics, and specific contexts where each approach, or a combination of both, may be most effective. By examining the relative strengths and weaknesses of different tools, comparative studies will guide providers and researchers in selecting the most appropriate tools based on specific contexts and populations.

Implications of findings for new screening tools

Our scoping review reveals a critical need for practical, contemporary, multidimensional screening tools administered by primary care providers that effectively assess mental health and development in young children while guiding the next steps in clinical care. The following sections outline key considerations for developing, implementing, and evaluating such tools, organized around core quality metrics, implementation factors, and practical utility [61, 127–129].

Psychometric requirements

Future screening tools must meet rigorous psychometric standards through several key criteria [130]. These tools should maintain strong reliability, such as Cronbach's alpha coefficients of 0.80 or higher across administrators and time points [131]. Validation studies must employ robust methodology across diverse populations to ensure broad generalizability [132, 133]. These instruments should exhibit strong predictive and concurrent validity when compared to established gold standard assessments [134, 135]. Furthermore, to ensure accurate identification of cases while minimizing false positives, these tools should maintain balanced sensitivity and specificity thresholds exceeding 0.70 [9, 124].

Implementation and clinical integration

Successful integration into clinical practice demands careful consideration of implementation factors [38, 69, 129]. New screening tools must be efficient in administration time, resource utilization, and cost [136]. Additionally, these tools should integrate seamlessly into existing clinical workflows while aligning with current billing and reimbursement systems. Furthermore, developers must ensure sustainable implementation by providing ongoing support, maintenance, and evaluation protocols suitable for primary care settings.

Clinical utility and decision support

Screening tools must demonstrate clear value in supporting healthcare decisions. Effective tools should yield actionable results that directly guide referrals when required and inform subsequent assessment and intervention planning. These tools must offer advantages and protocols distinct from existing developmental monitoring or diagnostic approaches, providing unique clinical insights. Furthermore, they should align with established practice guidelines and available services to ensure meaningful integration into care pathways.

Accessibility and cultural adaptations

Screening tools must be accessible and user-friendly to facilitate widespread adoption. Content should be clear and comprehensible for both providers and patients, ensuring accurate assessment. Tools must demonstrate cultural sensitivity through appropriate language options and attention to cultural nuances across diverse populations. Additionally, they should offer flexible administration formats accomodating provider, parent, and professional completion (including allied health professionals and early childhood educators) to enhance utility across clinical and community settings [137].

Provider-centered design

Screening tools must align with clinical realities and provider workflows. Current 20-minute administration times often exceed the practical constraints of busy clinical settings, and briefer tools are required. Tools should also minimize training demands by leveraging existing provider competencies while offering clear protocols for introduction, administration, scoring, and follow-up that accommodate competing clinical demands [129].

Strengths and limitations of this review

Strengths

A strength of this review includes using a comprehensive database search strategy developed by a Research Librarian and snowball searching to ensure relevant articles were captured. Two independent reviewers reviewed articles to reduce selection bias. We emailed the

corresponding authors of articles when additional information or clarification was required. Although tools had to be available in English and/or French to be included in this review, we identified which were adapted for use in other cultures, languages, or countries and may have been used outside primary care settings.

Limitations

This scoping review focused on published, brief, multidimensional screening tools for young children's mental health and development designed for primary care settings, representing a subset of available instruments. We excluded tools designed for other professionals (e.g., emergency department staff, educators) even if potentially suitable for primary care use. Clinical implementation may differ from research settings, as providers might adapt administration protocols. The varying depth and quality of tool descriptions in the literature may have affected our interpretation. Additionally, methodological quality assessment was beyond our scope, and not all article data could be synthesized. Single screening tools may not capture all domains of mental health and development. Tools focusing exclusively on specific aspects (e.g., autism, language skills) or addressing only mental health or development separately were excluded [19]. Despite our comprehensive search strategy, we may have missed eligible tools, particularly those not published in peer-reviewed articles or those available only in languages other than English or French.

Conclusions

This scoping review identified three brief, multidimensional screening tools for assessing mental health and development in children up to 6 years of age, described in six published articles. While these provider-administered tools shared core elements, their specific protocols, items, domains, and psychometric properties differed. Current research gaps include limited recent data and use and insufficient understanding of practical implementation, cost-effectiveness, and adaptability across different settings. Future studies should investigate how effectively these screening tools can be administered and integrated into clinical practice, compare their utility against parent-completed assessments, and assess their economic impact on healthcare systems. The field urgently needs new screening tools that maintain strong psychometric properties while addressing current limitations. Such tools should seamlessly integrate into clinical workflows, offer streamlined administration, and provide clear guidance for clinical decisions, all while ensuring cultural appropriateness and broad accessibility. Both newly developed tools and studies are essential for improving early identification and intervention for young children's mental health and developmental concerns.

Abbreviations

DQ	Developmental quotient
DSI	Developmental Screening Inventory
JBI	Formerly known as the Joanna Briggs Institute
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
WSST	Woodside System Screening Technique

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12875-025-02869-z.

Supplementary Material 1: Additional File 1_Abbreviated electronic database search strategy.docx

Supplementary Material 2: Additional File 2_A comprehensive overview of the characteristics of included a.docx

Supplementary Material 3: Additional File 3_Psychometric properties of the screening tools.docx

Acknowledgements

We thank Karine Fournier (KF), MIS (Research Librarian, University of Ottawa Library), for generating and conducting the search strategy and reviewing earlier versions of the manuscript, Marie-Cécile Domecq (MCD), MLIS (Research Librarian, University of Ottawa Library) for her peer review of the MEDLINE search strategy, Allie Russell (AR; former undergraduate student) and Helen Valkanas (HV; former Research Coordinator) for their independent reviews of the literature, and Dr. Hilary Myron (Pediatrician, Monfort Hospital and CHEO in Ottawa, Ontario, Canada) for her review and feedback on the manuscript.

Author contributions

We confirm that all named authors meet the criteria for authorship. ADR, CP, PC, and MC conceived and designed the study. ADR, CP, PC, KB, and MC participated in the acquisition, analysis, and interpretation of the data. All authors drafted the manuscript and/or substantively revised it. All authors approved the submitted version and agreed to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Data availability

Data is provided within the manuscript or additional files. Further details about the included articles or screening tools can be found in the original research articles cited in this review.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Knowledge Institute on Child and Youth Mental Health and Addictions, Ottawa, ON, Canada
[2]CHEO (Children's Hospital of Eastern Ontario) Research Institute, 401 Smyth Road, K1H 8L1 Ottawa, ON, Canada
[3]Department of Psychology, Carleton University, Ottawa, ON, Canada
[4]CHEO, Ottawa, ON, Canada
[5]Department of Psychiatry, University of Ottawa, Ottawa, ON, Canada
[6]School of Psychology, University of Ottawa, Ottawa, ON, Canada

Received: 16 July 2024 / Accepted: 1 May 2025

Published online: 09 June 2025

References

1. Glascoe FP. Early detection of developmental and behavioral problems. Pediatr Rev. 2000;21:272–80.
2. Lavigne JV, Binns HJ, Christoffel KK, Rosenbaum D, Arend R, Smith K, et al. Behavioral and emotional problems among preschool children in pediatric primary care: prevalence and pediatricians' recognition. Pediatrics. 1993;91:649–55.
3. Costello EJ, Burns BJ, Costello AJ, Edelbrock C, Dulcan M, Brent D. Service utilization and psychiatric diagnosis in pediatric primary care: the role of the gatekeeper. Pediatrics. 1988;82:435–41.
4. Sheldrick RC, Merchant S, Perrin EC. Identification of developmental-behavioral problems in primary care: A systematic review. Pediatrics. 2011;128:356–63.
5. Center on the Developing Child at Harvard University. The foundations of lifelong health are built in early childhood. [Internet]. 2010. Available from: https://developingchild.harvard.edu/wp-content/uploads/2010/05/Foundations-of-Lifelong-Health.pdf
6. Zeanah CH. Handbook of infant mental health. Guilford; 2018.
7. Kemper KJ. Self-administered questionnaire for structured psychosocial screening in pediatrics. Pediatrics. 1992;89:433–6.
8. Bekman S, St. John-Larkin C, Paul JJ, Millar A, Frankel K. Update on screening, referring, and treating the behavioral, social, and mental health problems of very young children. Curr Treat Options Pediatr. 2017;3:15–31.
9. Bright Futures Steering Committee, Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. Pediatrics. 2006;118:405–20.
10. Kieling C, Baker-Henningham H, Belfer M, Conti G, Ertem I, Omigbodun O, et al. Child and adolescent mental health worldwide: evidence for action. Lancet. 2011;378:1515–25.
11. Simpson TE, Condon E, Price RM, Finch BK, Sadler LS, Ordway MR. Demystifying infant mental health: what the primary care provider needs to know. J Pediatr Health Care. 2016;30:38–48.
12. American Academy of Pediatrics. Bright Futures: Guidelines for health supervision of infants, children, and adolescents. 4th edition. American Academy of Pediatrics; 2017.
13. Canadian Task Force on Preventive Health Care. Recommendations on screening for developmental delay. CMAJ. 2016;188:579–87.
14. Foy JM, Green CM, Earls MF, Lavin A, Askew GL, Baum R et al. Mental health competencies for pediatric practice. Pediatrics 2019;144e20192757.
15. Lipkin PH, Macias MM, Norwood KW, Brei TJ, Davidson LF, Davis BE et al. Promoting optimal development: identifying infants and young children with developmental disorders through developmental surveillance and screening. Pediatrics 2020;145:e20193449.
16. The Royal Australian College of General Practitioners. Guidelines for preventive activities in general practice [Internet]. 9th edition. East Melbourne, Vic: RACGP. 2018. Available from: https://www.racgp.org.au/FSDEDEV/media/documents/Clinical%20Resources/Guidelines/Red%20Book/Guidelines-for-preventive-activities-in-general-practice.pdf
17. US Preventive Services Task Force. Screening for anxiety in children and adolescents: US preventive services task force recommendation statement. JAMA. 2022;328:1438–44.
18. US Preventive Services Task Force. Screening for depression and suicide risk in children and adolescents: US preventive services task force recommendation statement. JAMA. 2022;328:1534–42.
19. Weitzman C, Wegner L, Section on Developmental and Behavioral Pediatrics C on PA of C and FH Council on Early Childhood, and Society for Developmental and Behavioral Pediatrics, Blum NJ, Macias MM, Bauer NS, et al. Promoting optimal development: screening for behavioral and emotional problems. Pediatrics. 2015;135:384–95.
20. Hackell JM, Abularrage JJ, Almendarez YM, Arauz Boudreau AD, Berhane AM, Cantrell PE et al. Recommendations for preventive pediatric health care. Pediatrics 2021;147:e2020049776..
21. Muldoon LK, Hogg WE, Levitt M. Primary care (PC) and primary health care (PHC) what is the difference?? Can. J Public Health. 2006;97:409–11.
22. World Health Organization. Declaration of alma-ata. World Health Organization. Regional Office for Europe; 1978.
23. Starfield B. Primary care: balancing health needs, services and technology. 2nd edition. New York: Oxford University Press; 1998.
24. Donaldson MS, Yordy KD, Lohr KN, Vanselow NA. Primary care: America's health in a new era. 1996.
25. World Health Organization, United Nations Children's Fund (UNICEF). A vision for primary health care in the 21st century: towards universal health coverage and the Sustainable Development Goals [Internet]. Geneva: World Health Organization. 2018. Available from: https://iris.who.int/handle/10665/328065
26. Gleason MM. Infant mental health in primary care. In: Zeanah CH, editor. Handbook of infant mental health. New York, NY: The Guilford Press; 2019. pp. 568–80.
27. Anderson Rothman A, Wagner EH, Showstack JA. Chronic illness management: what is the role of primary care? Ann Intern Med. 2003;138:256–61.
28. Simonian SJ. Screening and identification in pediatric primary care. Behav Modif. 2006;30:114–31.
29. Bright Futures /American Academy of Pediatrics. Recommendations for Preventive Pediatric Health Care—Periodicity Schedule [Internet]. 2024;Available from: https://www.aap.org/en/practice-management/care-delivery-approaches/periodicity-schedule/
30. Lavigne JV, Gibbons RD, Arend R, Rosenbaum D, Binns HJ, Christoffel KK. Rational service planning in pediatric primary care: continuity and change in psychopathology among children enrolled in pediatric practices. J Pediatr Psychol. 1999;24:393–403.
31. Bricker D, Davis MS, Squires J. Mental health screening in young children. Infants Young Child. 2004;17:129–44.
32. Carter AS, Briggs-Gowan MJ, Davis NO. Assessment of young children's social-emotional development and psychopathology: recent advances and recommendations for practice. J Child Psychol Psychiatry. 2004;45:109–34.
33. Rourke L, Leduc D. Improving the odds for effective developmental surveillance. Paediatr Child Health. 2012;17:539–40.
34. Committee on Children with Disabilities. Developmental surveillance and screening of infants and young children. Pediatrics. 2001;108:192–5.
35. Bobrowska A, Murton M, Seedat F, Visintin C, Mackie A, Steele R et al. Targeted screening in the UK: A narrow concept with broad application. Lancet Reg Heal. 2022;1 6:100353.
36. Pizur-Barnekow K, Erickson S, Johnston M, Bass T, Lucinski L, Bleuel D. Early identification of developmental delays through surveillance, screening, and diagnostic evaluation. Infants Young Child. 2010;23:323–30.
37. Ammitzbøll J, Holstein BEBE, Wilms L, Andersen A, Skovgaard AM, Ammitzboll J, et al. A new measure for infant mental health screening: development and initial validation. BMC Pediatr. 2016;16:197.
38. Rydz D, Shevell MI, Majnemer A, Oskoui M. Topical review: developmental screening. J Child Neurol. 2005;20:4–21.
39. Schonwald A, Huntington N, Chan E, Risko W, Bridgemohan C. Routine developmental screening implemented in urban primary care settings: more evidence of feasibility and effectiveness. Pediatrics. 2009;123:660–8.
40. Sices L. Developmental screening in primary care: The effectiveness of current practice and recommendations for improvement. 2007.
41. Godoy L, Carter AS. Identifying and addressing mental health risks and problems in primary care pediatric settings: A model to promote developmental and cultural competence. Am J Orthopsychiatry. 2013;83:73–88.
42. Hayutin LG, Reed-Knight B, Blount RL, Lewis J, McCormick ML. Increasing Parent–Pediatrician communication about children's psychosocial problems. J Pediatr Psychol. 2009;34:1155–64.
43. Sheldrick RC, Marakovitz S, Garfinkel D, Carter AS, Perrin EC. Comparative accuracy of developmental screening questionnaires. JAMA Pediatr. 2020;174:366–74.
44. Wissow LS, van Ginneken N, Chandna J, Rahman A. Integrating children's mental health into primary care. Pediatr Clin North Am. 2016;63:97–113.
45. King TM, Tandon SD, Macias MM, Healy JA, Duncan PM, Swigonski NL, et al. Implementing developmental screening and referrals: lessons learned from a National project. Pediatrics. 2010;125:350–60.
46. Sand N, Silverstein M, Glascoe FP, Gupta VB, Tonniges TP, O'Connor KG. Pediatricians' reported practices regarding developmental screening: do guidelines work? Do They Help?? Pediatr. 2005;116:174–9.
47. Radecki L, Sand-Loud N, O'Connor KG, Sharp S, Olson LM. Trends in the use of standardized tools for developmental screening in early childhood: 2002–2009. Pediatrics. 2011;128:14–9.
48. Van Cleave J, Kuhlthau KA, Bloom S, Newacheck PW, Nozzolillo AA, Homer CJ, et al. Interventions to improve screening and follow-up in primary care: a systematic review of the evidence. Acad Pediatr. 2012;12:269–82.

49. Sices L, Feudtner C, McLaughlin J, Drotar D, Williams M. How do primary care physicians identify young children with developmental delays? A National survey. J Dev Behav Pediatr. 2003;24:409–17.
50. Godoy L, Gordon S, Druskin L, Long M, Kelly KP, Beers L. Pediatric provider experiences with implementation of routine mental health screening. J Dev Behav Pediatr. 2021;4 2:32-40.
51. Health AA, of C and APC on HCA and ETF on M. Improving mental health services in primary care: reducing administrative and financial barriers to access and collaboration. Pediatrics. 2009;123:1248–51.
52. Godoy L, Carter AS, Silver RB, Dickstein S, Seifer R. Infants and toddlers left behind: mental health screening and consultation in primary care. J Dev Behav Pediatr. 2014;35:334–43.
53. Foy JM, Health AA. Of PTF on M. Enhancing pediatric mental health care: report from the American academy of pediatrics task force on mental health. Introduction Pediatr. 2010;125:S69.
54. Cappelli M, Leon SL. Paving the path to connected care: strengthening the interface between primary care and community-based child and youth mental health services. Ottawa, ON: Ontario Centre of Excellence for Child and Youth Mental Health; 2017.
55. Tanner JL, Stein MT, Olson LM, Frintner MP, Radecki L. Reflections on well-child care practice: a National study of pediatric clinicians. Pediatrics. 2009;124:849–57.
56. Morelli DL, Pati S, Butler A, Blum NJ, Gerdes M, Pinto-Martin J, et al. Challenges to implementation of developmental screening in urban primary care: a mixed methods study. BMC Pediatr. 2014;14:1–11.
57. Wissow LS, Brown J, Fothergill KE, Gadomski A, Hacker K, Salmon P, et al. Universal mental health screening in pediatric primary care: A systematic review. J Am Acad Child Adolesc Psychiatry. 2013;52:1134–e114723.
58. Rosenberg SA, Zhang D, Robinson CC. Prevalence of developmental delays and participation in early intervention services for young children. Pediatrics. 2008;121:e1503–9.
59. Ayres CG, Griffith HM. Perceived barriers to and facilitators of the implementation of priority clinical preventive services guidelines. Am J Manag Care 2007;13 150-155.
60. 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. 2024;67:286-306..
61. Feeney-Kettler KA, Kratochwill TR, Kaiser AP, Hemmeter ML, Kettler RJ. Screening young children's risk for mental health problems: A review of four measures. Assess Eff Interv. 2010;35:218–30.
62. Lavigne JV, Meyers KM, Feldman M. Systematic review: classification accuracy of behavioral screening measures for use in integrated primary care settings. J Pediatr Psychol. 2016;41:1091–109.
63. Marlow M, Servili C, Tomlinson M. A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low-and middle-income countries. Autism Res. 2019;12:176–99.
64. Rah SS, Jung M, Lee K, Kang H, Jang S, Park J et al. Systematic Review and Meta-Analysis: Real-World Accuracy of Children's Developmental Screening Tests. J. Am. Acad. Child Adolesc. Psychiatry [Internet]. 2022;Available from: https://www.sciencedirect.com/science/article/pii/S0890856722020330
65. Squires J, Bricker D. Ages and Stages Questionnaires. 3rd edition. Baltimore, MD: Paul H. Brookes Publishing Co.; 2009.
66. Glascoe F. Parents' Evaluation of Developmental Status - Revised (PEDS-R) [Internet]. Nolensville, Tennessee: PEDStest.com, LLC; 2023. Available from: www.pedstest.com.
67. Pontoppidan M, Niss NK, Pejtersen JH, Julian MM, Væver MS. Parent report measures of infant and toddler social-emotional development: a systematic review. Fam Pract. 2017;34:127–37.
68. Aylward GP. Developmental screening and assessment: what are we thinking? J Dev Behav Pediatr. 2009;30:169–73.
69. Glascoe FP. Screening for developmental and behavioral problems. Ment Retard Dev Disabil Res Rev. 2005;11:173–9.
70. Penumalee L, Lambert JO, Gonzalez M, Gray M, Partani E, Wilson C, et al. Why do they want to know?? A qualitative assessment of caregiver perspectives on social drivers of health screening in pediatric primary care. Acad Pediatr. 2023;23:329–35.
71. Schneider M, Mehari K, Langhinrichsen-Rohling J. What caregivers want: preferences for behavioral health screening implementation procedures in pediatric primary care. J Clin Psychol Med Settings. 2021;28:562–74.
72. Glascoe FP, Leew S. Parenting behaviors, perceptions, and psychosocial risk: impacts on young children's development. Pediatrics. 2010;125:313–9.
73. Agazzi H, Shaffer-Hudkins EJ, Armstrong KH, Hayford H, Agazzi H, Shaffer-Hudkins EJ et al. Best practices in assessment and intervention with infants and toddlers. Promot. Posit. Behav. Outcomes infants toddlers. Evidence Based Guide Early Interv. 2020;21–33.
74. Crossing the Quality Chasm. A New Health System for the 21st Century. Washington (DC): 2001.
75. Halfon N, Stevens GD, Larson K, Olson LM. Duration of a well-child visit: association with content, family-centeredness, and satisfaction. Pediatrics. 2011;128:657–64.
76. Norlin C, Crawford MA, Bell CT, Sheng X, Stein MT. Delivery of Well-Child care: A look inside the door. Acad Pediatr. 2011;11:18–26.
77. Halfon N, Inkelas M, Mistry R, Olson LM. Satisfaction with health care for young children. Pediatrics. 2004;113:1965–72.
78. Canadian Paediatric Society. You and your child's doctor [Internet]. Caringforkids.cps.ca2018;Available from: https://caringforkids.cps.ca/handouts/health-conditions-and-treatments/your_childs_doctor
79. Johnson CP. Using developmental and behavioral screening tests. Pediatr Rev. 2000;21:255–6.
80. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol Theory Pract. 2005;8:19–32.
81. Teare G, Taks M. Extending the scoping review framework: a guide for interdisciplinary researchers. Int J Soc Res Methodol. 2020;23:311–5.
82. Peters M, Godfrey C, McInerney P, Soares C, Hanan K, Parker D. The Joanna Briggs Institute Reviewers' Manual 2015: Methodology for JBI Scoping Reviews. 2015.
83. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth. 2020;18:2119–26.
84. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg. 2021;88:105906.
85. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169:467–73.
86. Radomski AD, Polihronis C, Cloutier P, Beaudin K, Cappelli M. Scoping Review of Multidimensional Mental Health and Developmental Screening Tools Administered by Primary Care Providers for Children Up to Six Years of Age [Internet]. 2023;Available from: https://archive.org/details/osf-registrations-5fby9-v1
87. Rourke JT, Rourke LL. Well baby visits: screening and health promotion. Can Fam Physician Med Fam Can. 1985;31:997–1002.
88. Li P, Rowan-Legg A, Kwok B, Bayoumi I, Arulthas S, Tedone E, et al. 2020 Edition of the Rourke baby record. Can Fam Physician. 2021;67:488.
89. Cairney J, Clinton J, Veldhuizen S, Rodriguez C, Missiuna C, Wade T, et al. Evaluation of the revised nipissing district developmental screening (NDDS) tool for use in general population samples of infants and children. BMC Pediatr. 2016;16:42–42.
90. Bayley N. Bayley scales of infant and toddler development - Third edition. San Antonio, TX: Harcourt Assessment, Psych. Corporation; 2006.
91. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol. 2016;75:40–6.
92. Covidence systematic review software [Internet]. Available from: www.covidence.org.
93. Zotero. [Internet]. Available from: https://www.zotero.org/
94. Clinton J, Feller A, Williams R. The importance of infant mental health. Paediatr Child Health. 2016;21:239–41.
95. Bornstein MH, Lansford JE. Assessing early childhood development. Handb Early Child Dev Res its Impact Glob Policy 2013;351–70.
96. Alfonso VC, Engler JR, Lepore JC. Assessing and evaluating young children: Developmental domains and methods. 2020.
97. Van Hemel SB, Snow CE. Early childhood assessment: why, what, and how. National Academies; 2008.
98. Liljenquist K, Hurst R, Guerra LS, Szilagyi PG, Fiscella K, Porras-Javier L, et al. Time spent at well-child care visits for English-and Spanish-speaking parents. Acad Pediatr. 2023;23:359–62.
99. Irving G, Neves AL, Dambha-Miller H, Oishi A, Tagashira H, Verho A, et al. International variations in primary care physician consultation time: a systematic review of 67 countries. BMJ Open. 2017;7:e017902.

100. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.
101. Sun S. Meta-analysis of Cohen's kappa. Health Serv Outcomes Res Methodol. 2011;11:145–63.
102. Pollock D, Peters MD, Khalil H, McInerney P, Alexander L, Tricco AC, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth. 2023;21:520–32.
103. Peters M, Godfrey C, McInerney P, Munn Z, Tricco A, Khalil H. Scoping Reviews (2020 version) [Internet]. In: JBI Manual for Evidence Synthesis. JBI; 2020. Available from: https://synthesismanual.jbi.global
104. Glascoe FP. The brigance infant and toddler screen: standardization and validation. J Dev Behav Pediatr. 2002;23:145–50.
105. Knobloch H, Pasamanick B, Sherard ES. A developmental screening inventory for infants. Pediatrics. 1966;38:viii–1.
106. Aina OF, Morakinyo O. The validation of developmental screening inventory (DSI) on Nigerian children. J Trop Pediatr. 2001;47:323–8.
107. Aina OF, Morakinyo O. Normative data on mental and motor development in Nigerian children. West Afr J Med. 2005;24:151–6.
108. Eu B. Evaluation of a developmental screening system for use by child health nurses. Arch Dis Child. 1986;61:34.
109. Brigance A. Brigance preschool screen. North Billerica, MA: Curriculum Associates, Inc.; 1985.
110. Brigance A, Brigance K. 1 Screen. North Billerica. MA: Curriculum Associates, Inc.; 1985.
111. Brigance A. Brigance early preschool screen. North Billerica, MA: Curriculum Associates, Inc.; 1990.
112. Brigance A. Brigance inventory of early development. North Billerica, MA: Curriculum Associates, Inc.; 1991.
113. Bobo C. K & 1 Screen as a Predictor for Early Identification of Students At Risk for Later Special Education Placement. Dissertation Abstracts International; 1992. page 1475.
114. Knobloch H, Gesell AL, Amatruda CS, Pasamanick B. Gesell and Amatruda's developmental diagnosis: the evaluation and management of normal and abnormal neuropsychologic development in infancy and early childhood. Harper & Row; 1974.
115. Gesell AL, Amatruda CS. Developmental diagnosis. 2nd ed. New York: Paul B. Hoeber, Inc.; 1954.
116. Knobloch H, Rider R, Pasamanick B, Harper P. An evaluation of a questionnaire on infant development. Am J Public Health Nations Health. 1955;45:1309–20.
117. Barber J, Boothman R, Stanfield J. A new visual chart for pre-school developmental screening. Health Bull (Edinb). 1976;34:80–91.
118. Bayley N. Manual for the Bayley scales of infant development. Psychological Corporation; 1969.
119. Frankenburg WK, Goldstein AD, Camp BW. The revised Denver developmental screening test: its accuracy as a screening instrument. J Pediatr. 1971;79:988–95.
120. Griffiths R. The abilities of babies: A study in mental measurement. London: University of London Press Lrd.; 1954.
121. Griffiths R. The abilities of young children: A comprehensive system of mental measurement for the first eight years of life. London: Child Development Research Centre; 1970.
122. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med. 2012;22:276–82.
123. White RF, Braun JM, Kopylev L, Segal D, Sibrizzi CA, Lindahl AJ, et al. NIEHS report on evaluating features and application of neurodevelopmental tests in epidemiological studies. NIEHS; 2022. Report 01 [Internet].
124. Trevethan R. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. Front Public Health. 2017;5:308890.
125. Lipkin PH, Macias MM, Baer Chen B, Coury D, Gottschlich EA, Hyman SL et al. Trends in pediatricians' developmental screening: 2002–2016. Pediatrics 2020;14(5):e20190851. https://doi.org/10.1542/peds.2019-0851
126. Iragorri N, Spackman E. Assessing the value of screening tools: reviewing the challenges and opportunities of cost-effectiveness analysis. Public Health Rev. 2018;39:17.
127. Foy JM, Kelleher KJ, Laraque D, American Academy of Pediatrics Task Force on Mental Health. Enhancing pediatric mental health care: strategies for Preparing a primary care practice. Pediatrics. 2010;125:S87–108.
128. Glover TA, Albers CA. Considerations for evaluating universal screening assessments. J Sch Psychol. 2007;45:117–35.
129. Drotar D, Stancin T, Dworkin P. Pediatric developmental screening: understanding and selecting screening instruments. Citeseer; 2008.
130. McKechnie D, Fisher MJ. Considerations when examining the psychometric properties of measurement instruments used in health. TheAustralian J Adv Nurs. 2022;39:36–47.
131. Nunnally J, Bernstein J. Psychometric Theory, 3rd EditionMcGraw-Hill. N. Y. NY. 1994.
132. Bell N, Gorber SC, Tonelli M, Pottie K, Singh H, Joffres M, et al. From ABCs to GRADE: Canadian task force on preventive health Care's new rating system for clinical practice guidelines. Can Fam Physician. 2013;59:1282–9.
133. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336:924–6.
134. Taherdoost H. Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. Test Valid. Quest. Res. August 10 2016 2016.
135. Alavi M, Thapa DK, Biros E, Cleary M. Considerations in the use of concurrent or predictive validity in clinical measurement. Nurs Health Sci. 2023;25:183.
136. Fischer VJ, Morris J, Martines J. Developmental screening tools: feasibility of use at primary healthcare level in low-and middle-income settings. J Health Popul Nutr. 2014;32:314–314.
137. Ertem IO, Dogan DG, Gok CG, Kizilates SU, Caliskan A, Atay G et al. A guide for monitoring child development in low-and middle-income countries. Pediatrics 2008;121:e581-9.

Publisher's note