New trends in instruments for child development screening in Brazil: a systematic review

Karolina Alves de Albuquerque¹, Ana Cristina Barros da Cunha²

Abstract

Introduction: Screening instruments are widely used to monitor child development. The accurate use of standardized tools is an indispensable condition for clinical practice and research aimed at detecting developmental risks and other problems in children.

Objective: The objective of this systematic review was to analyze the use of standardized tools for child development screening used in studies with Brazilian children.

Methods: Two independent researchers selected references in English and Portuguese from five databases through which they searched for studies that used screening tests to assess the development of Brazilian children. All articles were read to determine the main objective, design, target population, the type of screening test, and the purpose of using the test with Brazilian children.

Results: Among the 27 papers analyzed, most of them was observational studies conducted with children up to six years of age, with the main objective to screen development delays and analyze associations between risks and child development. Four instruments were identified: Denver Developmental Screening Test II, Ages and Stages Questionnaire, Bayley Scales of Infant and Toddler Development Screening Test, and Battelle Developmental Inventory Screening Test. Three of these tests have been validated for use in Brazil.

Conclusion: This review suggests that the screening instruments have been used in research for different purposes, such as in the diagnosis of developmental problems, and sometimes inappropriately. Furthermore, studies to validate measures for screening and assessing the development of Brazilian children are still scarce and, therefore, deserve more attention.

Keywords: child development, screening instruments, surveillance, development evaluation, systematic review.
The monitoring of child development is a broad, continuous, and indispensable process for the health and development of children\(^1\), with particular importance given the most recent public policies for child health care\(^1\). In this process, the most frequently used screening measure by Brazilian health professionals to assess the changes in child development is an informal clinical assessment\(^4\). However, the clinical judgment based on this assessment detects only 30% of children suspected of being at-risk for development problems. In comparison, standardized screening instruments have a 70% to 90% sensitivity and specificity when it comes to identifying these children\(^7\).

Standardized assessment instruments have been used for different purposes. Screening tests, for example, are used to identify children at risk for developmental delays and, therefore, need to be adequate for specific assessment\(^7\). In fact, they must be quick to administer and comprehensive enough to indicate problems in different development domains, such as motor, cognitive and socioemotional. Standardized screening instruments are the best tools for the early identification of children at risk for delays\(^1\). However, they are not sufficient to describe or diagnose specific developmental changes due to their features\(^7\).

Currently, most of the child developmental screening tests available have been developed in North America. To be applied in Brazil, they need to undergo a process of translation, adaptation, and verification of their psychometric properties to be a good tool to evaluate Brazilian children\(^7\).\(^8\). Previous review studies of child development assessment instruments pointed out the lack of validated tests in Brazil\(^9\)\(^-\)\(^10\). It should also be considered that these studies reviewed different child assessment measures and not only screening tests. Thus, those studies discuss the standardized child development screening tests available in Brazil and their effectiveness for the early detection of problems in a large process of developmental monitoring. Based on this finding, the guiding question for this systematic review was: “What standardized child development screening tests are most frequently used with Brazilian children, and how they have been used in research?”\(^7\). Thus, the objective of this review was to examine the use of standardized child development screening tests adopted in studies with Brazilian children.

### METHODS

This is a systematic review study conducted according to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses - PRISMA. The references were searched in five indexed databases: PubMed, Virtual Health Library (VHL), LILACS, IBECS, and SciELO. Two independent researchers carried out the search and selection process from April to May 2018 and updated in February 2020. The following keywords and Boolean operators were used: “child development” OR “neuropsychomotor development” OR “infant development”, AND screening OR surveillance, AND “Brazil”, and their respective terms in Portuguese. All divergences and disagreements in the search and selection of papers were solved by the consensus between independent researchers. The Mendeley Desktop\(^\circ\) software was used to organize the references.

References were retrieved following these criteria: a) empirical studies in English and Portuguese; b) studies published as a scientific paper in national or international journals in the last seven years (from 2014 to 2020); and c) only studies conducted with Brazilian children evaluated by a standardized child development screening test. For the purpose of this review, only screening tests with standardized procedures for application, correction and interpretation, rapid administration and that evaluated different development domains were considered\(^2\). References were excluded according to these criteria: a) using non-standard instruments or no screening tests; and b) not using a test to evaluate the child’s development.

After a complete reading of all articles, the following information was extracted: a) the main objective, research design, and target population of each study; b) the screening test used; and c) the purpose of the test used. Manuals, books, and sites about those tests were used to complement the information. Next, the general features of the tests in terms of age, the application and interpretation procedures, the sensitivity/specificity of the original version, and the adaptation/validation version available for use in Brazil were better described. Data was then organized in figures and tables, as presented below.

### RESULTS

The search processes and their results are described on Figure 1.
First, a total of 2139 references was searched. After applying the filters and removing duplicated references, 754 papers were examined by reading their titles. After reading the papers’ abstracts, 33 references were identified in accordance with the objective review. Then, all of them were completely read and six references were excluded because they used child evaluation instruments that were neither screening nor standardized instruments or used the instrument to verify the comprehension of adults about child development without testing the child.

In general, most of the references were observational transversal (n=18) or methodological (n=5) studies aimed at evaluating psychometric properties of the test. The observational studies aimed to verify development outcomes or associate the outcomes with bio psychosocial factors (n=16). Considering all the 27 articles retrieved, 77,943 Brazilian children under six years of age were evaluated. It is important to highlight that six studies included only children under two years of age11-16. For the most part, the target populations were children with normal development11,12,13,17-29 or children at-risk for delays30-36. Two studies included children with confirmed disability diagnoses, such as cerebral palsy37 and microcephaly13.

Four child development screening tests were identified on this review: 1) Denver Developmental Screening Test Revised – Denver II38; 2) Ages and Stages Questionnaire – ASQ39; 3) Bayley Scales of Infant and Toddler Development, Screening Test – Bayley-ST40; and 4) Battelle Developmental Inventory, Screening – BDIS41. The map of references that used these tests is presented in Figure 2.

A brief description of the tests in terms of content features, format, and psychometric properties of the original version is presented in Table 1.

In the revised studies, the child development screening tests were chosen and adopted for different purposes, as presented in Table 2.
DISCUSSION

Considering the relevance of using standardized and validated measures to monitor child development, this review sought to analyze the use of standardized development screening tests adopted in the studies with Brazilian children in order to discuss new tendencies in applying these tests in Brazil. The findings revealed that the studies had adopted four screening tests: Denver Test II, Ages and Stages Questionnaires (ASQ); Bayley Scales of Infant and Toddler Development, Screening Test (Bayley-ST) and the Battelle Developmental Inventory Screening (BDIS).

The choice of tests for clinic or research purposes should be based on the objectives of the professional or researcher, as well as the profile of the population to be evaluated and the psychometric properties of the test. In general, the studies used tests for screening child development to assess children in the age group corresponding to the tests’ target population, thus reaffirming the appropriateness of using them for the population age profile they evaluate. In addition, the analysis of the studies’ objectives underscored the fact that many of those studies used the screening tests according to their respective purposes, such as verifying the incidence of suspicion of development delays and searching for relationships between risk factors and developmental outcomes in the infant population. However, this review also revealed that some tests were used for objectives that exceeded the purpose of this type of assessment tool.

Screening tests are adequate for screening and identifying the incidence of suspicion of developmental delays, yet is not always considered. Part of the reviewed studies adopted the screening test to determine developmental delays in a specific population. Embora nenhum dos estudos revisados tenha buscado formalizar um diagnóstico, as conclusões podem ser interpretadas para classificação diagnóstica da criança e não como suspeita ou rastreio de atrasos de desenvolvimento.

Although none of the reviewed studies sought to formalize a diagnosis, the conclusions could be interpreted for the child’s diagnostic classification and not as a suspicion or screening for developmental delays. For example, Dell’Agnolo et al. concluded that three boys in their study evaluated by the Denver Test II were diagnosed with

Table 1: General features of the child development screening tests used on the studies with Brazilian children.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Age group</th>
<th>Application mode</th>
<th>Application time (min.)</th>
<th>Sensibility</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denver II</td>
<td>0 to 72</td>
<td>Direct observation and interview</td>
<td>20 – 30</td>
<td>0.56</td>
<td>0.80</td>
</tr>
<tr>
<td>ASQ</td>
<td>1 to 66</td>
<td>Interview</td>
<td>10 – 15</td>
<td>0.70 – 0.90</td>
<td>0.76 – 0.91</td>
</tr>
<tr>
<td>Bayley-ST</td>
<td>1 to 42</td>
<td>Direct observation</td>
<td>15 - 25</td>
<td>0.75 – 0.86</td>
<td>0.75 – 0.86</td>
</tr>
<tr>
<td>BDIS</td>
<td>0 to 95 months</td>
<td>Free and structured observation and interview</td>
<td>10 – 30</td>
<td>0.75 – 0.93</td>
<td>0.83 – 0.91</td>
</tr>
</tbody>
</table>

Legend: Denver II: Denver Screening Test II; ASQ: Ages and Stages Questionnaires; Bayley-ST: Bayley Scales of Infant and Toddler Development, Screening Test; BDIS: Battelle Developmental Inventory Screening Test.

Table 2: Purposes of the use of child development screening tests adopted in the studies with Brazilian children.

<table>
<thead>
<tr>
<th>Purpose of the instrument used</th>
<th>Instruments used (study references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>To search relationships between risk factors and developmental outcomes</td>
<td>Denver II (12,35); Bayley-ST (11,36); BDIS (17).</td>
</tr>
<tr>
<td>To screen developmental delays</td>
<td>Denver II (18,20,25,28,30); ASQ (13,26); Bayley (34).</td>
</tr>
<tr>
<td>To identify developmental delays</td>
<td>Denver II (33); ASQ (15); Bayley-ST (16); BDIS (14).</td>
</tr>
<tr>
<td>To describe developmental profile</td>
<td>Denver II (30,37); Bayley-ST (32).</td>
</tr>
<tr>
<td>To evaluate the effectiveness of an intervention</td>
<td>Denver II (31).</td>
</tr>
<tr>
<td>To analyze the psychometric properties of an instrument</td>
<td>Denver II (22); ASQ (19,23).</td>
</tr>
<tr>
<td>To evaluate the psychometric properties of another instrument (a gold-standard tool)</td>
<td>Denver II (21,24,29).</td>
</tr>
</tbody>
</table>

Denver II: Denver Screening Test II; ASQ: Ages and Stages Questionnaires; Bayley-ST: Bayley Scales of Infant and Toddler Development, Screening Test; BDIS: Battelle Developmental Inventory Screening Test.
delayed speech development. The use of screening tests as a measure to describe a child’s development profile is also a critical point. Screening tests are brief tools with few items and do not allow this type of description. They are designed to identify children who have been recommended for a complementary evaluation.

Among the screening tests adopted in the reviewed papers, the Denver Test II was the one most frequently used in 12 studies. Although there is no authorized adaptation or validation of this test for use with a Brazilian population, two researchers used it as a gold-standard in the validation of an assessment measure. This choice was made based on the popularity of the Denver II Test, as well as the scarcity of validated screening tests for use with Brazilian children. Moreover, the Denver II Test is a tool recommended by the Brazilian Society of Pediatrics and it is inexpensive and easy to apply. However, one must be cautious when interpreting and generalizing the results of the Denver Test because of its limitations in terms of validation in Brazil, which was not observed in the reviewed studies. Screening tests, such as the Denver Test, identify children at risk of developmental delays. Therefore, they are not suitable for assessing children with a confirmed diagnosis of delay or disabilities. Lamônica et al. used the Denver Test to evaluate children diagnosed with cerebral palsy, a health and neurodevelopmental condition that, by definition, severe delays in motor development. The Denver II Test has also been used to verify the effectiveness of a nutritional intervention in the cognitive performance of children, which is also a misuse of the screening test, which is not appropriate for evaluating the effectiveness of an intervention study. In addition, cognition is not one of the domains assessed by the Denver Test, which may result in bias in the analysis of research results.

It should be noted that the correct choice of an assessment measure is particularly important. Each test was developed for specific purposes and objectives to evaluate a population in specific developmental domains. Previous reviews of developmental assessment tools have pointed out the lack of validated tests for Brazil. In addition to the five methodological studies included in this review, three of the four screening tests are in the process of being validated for use with Brazilian children. However, the availability of these validated tests is still restricted in terms of being applied in Brazil. The Denver II Test, for example, has a Portuguese version which is marketed in Brazil, but it has not been standardized or validated for use with Brazilian children. Conversely, the ASQ has a Portuguese version that has been adapted for use in Brazilian daycare centers: the ASQ-BR. Although the psychometric studies with ASQ-BR found good indices of construct validity and excellent item-to-test correlations, it is not available for commercial use. Among the four studies that used the ASQ, three adopted ASQ-BR to improve the indices of internal consistency and validity.

Contrarily the Bayley Scales has a Portuguese version marketed in Brazil, which includes full screening versions, but it has not been adapted for use with Brazilian children. In 2016, a full version of Bayley Scales III was adapted in a first validation study proposed by Madaschi and collaborators, but it is also not being marketed. The full version of the Bayley III Scales showed high convergent validity and good internal consistency to evaluate children from 12 months, but low stability and test-retest reliability indices were also found. None of the articles included in this review cited the Bayley validation study and there are no psychometric studies of the screening version of Bayley Scales for Brazil.

Finally, the Battelle Developmental Inventory Screening (BDIS) was also used in the reviewed studies, which adopted a freely translated version of the Spanish version of BDIS, conducted by the authors, without adaptation or validation for use with Brazilian children. It is important to highlight that the BDIS was recently adapted for the Brazilian population by the authors of this review and the psychometric results were promising, indicating excellent validity indices, based on the internal structure, concurrent and convergent validity, internal consistency, and test-retest and examiners reliabilities.

Due to the importance of monitoring child development for delays and conducting an early intervention, the choice of a test for screening and identifying risks to child development must be made clearly and accurately. Tests that assess an ample age range allow for extensive analysis and a more reliable clinical application because they allow longitudinal monitoring by the same parameters. Knowing the developmental domains that the test evaluates is also important and should be consistent with the researcher’s purpose, as well as with clinical professionals in the Early Care field. Furthermore, this review contributes to the practice of monitoring and screening child development because the main tests for screening developmental delays are critically analyzed. Also, trends in the use of these tools are discussed, presenting a critical perspective that can help when choosing child assessment measures for use in the clinic and in research.

Finally, the results of this review indicate the need for more investment in appropriate assessment measures for the child population in Brazil. Although the instruments validated for this population are scarce, it is possible to researchers are making considerable effort to adapt and find psychometric evidence for Brazilian versions of screening tests. This commitment is valid and deserves recognition from the academic-scientific community, but financial incentive is mandatory for new validation studies. In addition, professionals should be careful when choosing and using adapted and validated instruments. Studies of cultural adaptation and assessment of psychometric properties are essential so the screening tools, already considered robust and sensitive internationally, can be used without restrictions in the assessment of the child population in Brazil.

**CONCLUSION**

This review identified four standardized tests for screening development that are used in research with Brazilian children. These tests do not yet have consistent psychometric validity for use in Brazil, which should be considered when choosing an assessment tool. The results...
of this review show that the screening tests are sometimes used inappropriately, mainly for purposes that go beyond their stated objectives, such as diagnosing or monitoring the evolution and progression of child development. Finally, our findings indicate that validation studies of screening tests beyond cross-cultural adaptations are necessary and deserve attention. This means supporting research aimed at validating more reliable and sensitive measures for assessing the development of Brazilian children.


Resumo

Introdução: Instrumentos de triagem são usados para monitoramento do desenvolvimento infantil. O uso acurado de ferramentas padronizadas é condição indispensável para a prática clínica e pesquisas que visam detectar risco de desenvolvimento e problemas em crianças.

Objetivo: O objetivo desta revisão sistemática foi analisar o uso de instrumentos padronizados de triagem do desenvolvimento infantil adotados em estudos com crianças brasileiras.

Método: Dois pesquisadores independentes selecionaram em cinco bases de dados referencias em Inglês e Português onde buscaram estudos que usaram testes de triagem para avaliação do desenvolvimento de crianças brasileiras. Todos os artigos foram lidos para analisar o objetivo principal, delineamento, população-alvo, o tipo de teste de triagem e o propósito de uso do teste com crianças brasileiras.

Resultados: Dentre os 27 artigos analisados, a maioria deles eram estudos observacionais conduzidos com crianças até seis anos de idade com objetivo principal de rastrear atrasos e analisar associações entre riscos e desenvolvimento. Quatro instrumentos foram identificados: Teste de Triagem do Desenvolvimento de Denver II; Ages and Stages Questionnaire; Bayley Scales of Infant and Toddler Development, Screening Test; e Battelle Developmental Inventory Screening Test. Três testes estão sendo validados para uso no Brasil.

Conclusão: Esta revisão sugere que os instrumentos de triagem têm sido usados nas pesquisas com diferentes finalidades, por vezes de forma apropriada ou incorreta, como por exemplo para diagnosticar problemas de desenvolvimento. Além disso, os estudos de validação de medidas para triagem e avaliação do desenvolvimento de crianças brasileiras ainda são escassos e, por isso, merecem atenção.

Palavras-chave: desenvolvimento infantil; instrumentos de triagem; vigilância; avaliação do desenvolvimento; revisão sistemática.